Optimal recovery of linear operators from information of random functions

K. Yu. Osipenko

Moscow State University,
Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow

Abstract

The paper concerns problems of the recovery of linear operators defined on sets of functions from information of these functions given with stochastic errors. The constructed optimal recovery methods, in general, do not use all the available information. As a consequence, optimal methods are obtained for recovering derivatives of functions from Sobolev classes by the information of their Fourier transforms given with stochastic errors. A similar problem is considered for solutions of the heat equation.

Keywords: optimal recovery, random functions, Fourier transform
2010 MSC: 41A65, 41A46, 49N30, 60G35

1. Introduction

There are several approaches to recovery problems from inaccurate information. One of them concerns the case when the error in the initial data is deterministic. Quite a lot of works are devoted to this case. The main results can be found in [8], [9], [10], [12], [14] and the literature cited there.

Another approach is related to the fact that the initial information is considered to be given with a random error. There are also many papers dedicated to this topic. The following are the closest to the setting under consideration: [11], [2], [1], [15], [16], [13], [4]. A distinctive specificity of this paper is that the information used here is not random vectors, but random functions. Moreover, we consider the set of various random functions not only with the Gaussian noise.

2. General setting

Denote by $\mathcal{W}$ the set of functions $x(\cdot) \in L_2(\mathbb{R})$ for which

$$\int_{\mathbb{R}} \nu(t)|x(t)|^2\, dt < \infty,$$

Email address: kosipenko@yahoo.com (K. Yu. Osipenko)

where $\nu(\cdot)$ is continuous and positive almost everywhere. Put

$$W = \left\{ x(\cdot) \in \mathcal{W} : \int_{\mathbb{R}} \nu(t)|x(t)|^2\,dt \le 1 \right\}.$$

Consider the problem of optimal recovery of the operator $\Lambda x(\cdot) = \mu(\cdot)x(\cdot)$ on the class W by functions $x(\cdot)$ given with random errors (we assume that $\mu(\cdot)$ is continuous and such that Λ maps $\mathcal{W}$ into $L_2(\mathbb{R})$). More precisely, for a fixed $\delta > 0$ and every $x(\cdot) \in W$ we consider the set of random functions

$$Y_\delta(x(\cdot)) = \{\, y_\xi(\cdot) \in L_2(\mathbb{R}) : \mathbb{M}y_\xi(\cdot) = x(\cdot),\ \operatorname{Var} y_\xi(\cdot) \le \delta^2 \text{ a.e.} \,\},$$

where $\mathbb{M}X$ is the expectation of X and $\operatorname{Var} X = \mathbb{M}|X - \mathbb{M}X|^2$ is the variance of X. We will also assume that the set of these random functions and the corresponding probability measures are such that they allow for a change in integration so that the equalities are valid

$$\mathbb{M}\int_{\mathbb{R}} p(t)y_\xi(t)\,dt = \int_{\mathbb{R}} p(t)\mathbb{M}y_\xi(t)\,dt, \quad p(t) \in L_2(\mathbb{R}), \tag{1}$$

and

$$\mathbb{M}\int_{\mathbb{R}} |y_\xi(t)|^2\,dt = \int_{\mathbb{R}} \mathbb{M}|y_\xi(t)|^2\,dt. \tag{2}$$

As recovery methods we consider all possible mappings $\varphi\colon L_2(\mathbb{R}) \to L_2(\mathbb{R})$. The error of a method φ is defined as

$$e(\Lambda, W, \delta, \varphi) = \left(\sup_{\substack{x(\cdot)\in W \\ y_\xi(\cdot)\in Y_\delta(x(\cdot))}} \mathbb{M}\left(\|\Lambda x(\cdot) - \varphi(y_\xi(\cdot))(\cdot)\|^2_{L_2(\mathbb{R})} \right) \right)^{1/2}.$$

The problem is to find the error of optimal recovery

$$E(\Lambda, W, \delta) = \inf_{\varphi\colon L_2(\mathbb{R})\to L_2(\mathbb{R})} e(\Lambda, W, \delta, \varphi) \tag{3}$$

and a method on which this infimum is attained which is called optimal.

A similar setting was studied in [11], but there, instead of the set $Y_\delta(x(\cdot))$, the set of random functions $y_\xi(\cdot) = x(\cdot) + \xi(\cdot)$, where $\xi(\cdot)$ is the Gaussian noise, was considered. In [11], an optimal method was found among linear methods and it was shown that it is asymptotically optimal. Note, that even in the simplest one-dimensional case the optimal method is nonlinear (see [12]). The main difference of our approach is that by expanding the set of admissible random functions (not limited to Gaussian noise) we were able to obtain an exact lower bound (this was the most difficult part of the proof).

We assume that $|\mu(\cdot)|$ and $\nu(\cdot)$ are even functions, $|\mu(t)| > 0$ almost everywhere, and $|\mu(\cdot)|/\sqrt{\nu(\cdot)}$ is a monotonically decreasing function on $\mathbb{R}_+ = [0, +\infty)$. Put

$$f(s) = \int_{|t|\le s} \left(\frac{\sqrt{\nu(s)}}{|\mu(s)|} \frac{|\mu(t)|}{\sqrt{\nu(t)}} - 1 \right) \nu(t)\,dt. \tag{4}$$

It is easy to check that $f(\cdot)$ is a monotonically increasing function. Assume that $f(s) \to +\infty$ as $s \to +\infty$. Then for any $\delta > 0$ the equation $f(s) = \delta^{-2}$ has a unique solution t_δ.

Theorem 1. *Let $|\mu(\cdot)|$ and $\nu(\cdot)$ be even functions, $|\mu(t)| > 0$ almost everywhere, and $|\mu(\cdot)|/\sqrt{\nu(\cdot)}$ is a monotonically decreasing function on $\mathbb{R}_+$. Assume that $f(s) \to +\infty$ as $s \to +\infty$. Then for all $\delta > 0$*

$$E(\Lambda, W, \delta) = \delta\left(\int_{|t|\le t_\delta} |\mu(t)|^2 \left(1 - \frac{\sqrt{\nu(t)}}{|\mu(t)|}\frac{|\mu(t_\delta)|}{\sqrt{\nu(t_\delta)}}\right) dt\right)^{1/2}, \tag{5}$$

where t_δ is the unique solution of the equation $f(s) = \delta^{-2}$. Moreover, the method

$$\varphi(y_\xi(\cdot))(t) = \left(1 - \frac{\sqrt{\nu(t)}}{|\mu(t)|}\frac{|\mu(t_\delta)|}{\sqrt{\nu(t_\delta)}}\right)_+ \mu(t) y_\xi(t)$$

is optimal ($a_+ = \max\{a, 0\}$).

Proof. 1. The lower bound. Consider the set $[-A, -\widetilde{a}] \cup [\widetilde{a}, A] \subset \mathbb{R}$, $0 < \widetilde{a} < A$. Let us divide it on $2N$ parts by the points $\pm x_j$, $j = 0, 1, \ldots, N$, where

$$x_j = \widetilde{a} + j\frac{A - \widetilde{a}}{N}.$$

Set $a = (a_1, \ldots, a_{2N})$,

$$x_a(t) = \begin{cases} a_{2j-1}, & t \in [x_{j-1}, x_j),\ j = 1, \ldots N, \\ a_{2N-1}, & t = A, \\ a_{2j}, & t \in [-x_j, -x_{j-1}),\ j = 1, \ldots, N, \\ a_0, & t = -\widetilde{a}, \\ 0, & t \notin [-A, -\widetilde{a}] \cup [\widetilde{a}, A]. \end{cases}$$

Let $\tau = (\tau_1, \ldots, \tau_{2N})$, $\tau_1 \ge \tau_2 \ge \ldots \ge \tau_{2N} > 0$, and $x_\tau(\cdot) \in W$. Put

$$B = \{\, x_a(\cdot) \in W : a_j = \pm\tau_j, \quad j = 1, \ldots, 2N \,\}.$$

Set

$$p_j = \frac{\delta^2}{\delta^2 + \tau_j^2}, \quad j = 1, \ldots, 2n.$$

Due to the monotony conditions of τ_j, we have

$$0 < p_1 \le \ldots \le p_{2n} < 1.$$

Any $x(\cdot) \in B$ can be written in the form

$$x(\cdot) = \sum_{j=1}^{2n} s_j(x)\tau_j e_j(\cdot),$$

where $s_j(x) \in \{-1, 1\}$, and

$$e_{2j-1}(\cdot) = \chi_{[x_{j-1}, x_j)}(\cdot), \ j = 1, \dots N-1, \quad e_N(\cdot) = \chi_{[x_{N-1}, x_N]}(\cdot)$$
$$e_2(\cdot) = \chi_{[-x_1, -x_0]}(\cdot), \quad e_{2j}(\cdot) = \chi_{[-x_j, -x_{j-1})}(\cdot), \ j = 2, \dots, N$$

($\chi_\Omega(\cdot)$ is the characteristic function of the set Ω).

For each $x(\cdot) \in B$ we define the distribution $\eta(x)(\cdot)$ in the following way:

$$\eta(x)(\cdot) = \begin{cases} 0, & \text{with probability } p_1, \\ \dfrac{s_1(x)\tau_1}{1-p_1} e_1(\cdot) & \text{with probability } p_2 - p_1, \\ \displaystyle\sum_{j=1}^{2} \frac{s_j(x)\tau_j}{1-p_j} e_j(\cdot), & \text{with probability } p_3 - p_2, \\ \dots\dots\dots\dots\dots & \dots\dots\dots\dots\dots\dots\dots\dots \\ \displaystyle\sum_{j=1}^{2n-1} \frac{s_j(x)\tau_j}{1-p_j} e_j(\cdot), & \text{with probability } p_{2n} - p_{2n-1}, \\ \displaystyle\sum_{j=1}^{2n} \frac{s_j(x)\tau_j}{1-p_j} e_j(\cdot), & \text{with probability } 1 - p_{2n}. \end{cases}$$

Thus, we have the following distribution for $\eta(x)(t)$ if $t \in [x_{j-1}, x_j)$:

$$\eta(x)(t) = \begin{cases} 0, & \text{with probability } p_{2j-1}, \\ \dfrac{s_{2j-1}(x)\tau_{2j-1}}{1-p_{2j-1}}, & \text{with probability } 1 - p_{2j-1}, \end{cases}$$
$$j = 1, \dots, N-1,$$

if $t \in [x_{N-1}, x_N]$:

$$\eta(x)(t) = \begin{cases} 0, & \text{with probability } p_{2N-1}, \\ \dfrac{s_{2N-1}(x)\tau_{2N-1}}{1-p_{2N-1}}, & \text{with probability } 1 - p_{2N-1}, \end{cases}$$

if $t \in [-x_1, -x_0]$:

$$\eta(x)(t) = \begin{cases} 0, & \text{with probability } p_2, \\ \dfrac{s_2(x)\tau_2}{1-p_2}, & \text{with probability } 1 - p_2, \end{cases}$$

and if $t \in [-x_j, -x_{j-1})$:

$$\eta(x)(t) = \begin{cases} 0, & \text{with probability } p_{2j}, \\ \dfrac{s_{2j}(x)\tau_{2j}}{1-p_{2j}}, & \text{with probability } 1 - p_{2j}, \end{cases} \quad j = 2, \dots, N,$$

if $t \notin [-A, -\widetilde{a}] \cup [\widetilde{a}, A]$: $\eta(x)(t) = 0$.

It is easy to verify that $\mathbb{M}\eta(x)(\cdot) = x(\cdot)$. Moreover,

$$\operatorname{Var}\eta(x)(t) = \begin{cases} \delta^2, & t \in [-A, -\widetilde{a}] \cup [\widetilde{a}, A], \\ 0, & t \notin [-A, -\widetilde{a}] \cup [\widetilde{a}, A]. \end{cases}$$

Let us verify that (1) and (2) hold for $\eta(x)(\cdot)$. It is sufficient to check the validity of these equalities for each interval (x_{j-1}, x_j), $(-x_j, -x_{j-1})$, $j = 1, \ldots, N$. Let $t \in (x_{j-1}, x_j)$. Then for $p(\cdot) \in L_2(\mathbb{R})$ we have

$$\int_{x_{j-1}}^{x_j} p(t)\eta(x)(t)\,dt = \begin{cases} 0, & \text{with probability } p_{2j-1}, \\ \displaystyle\int_{x_{j-1}}^{x_j} \frac{s_{2j-1}(x)\tau_{2j-1}}{1 - p_{2j-1}}\,dt, & \text{with probability } 1 - p_{2j-1}. \end{cases}$$

Thus,

$$\mathbb{M}\left(\int_{x_{j-1}}^{x_j} p(t)\eta(x)(t)\,dt\right) = \int_{x_{j-1}}^{x_j} p(t)x(t)\,dt.$$

On the other hand,

$$\int_{x_{j-1}}^{x_j} p(t)\mathbb{M}\eta(x)(t)\,dt = \int_{x_{j-1}}^{x_j} p(t)x(t)\,dt.$$

We have

$$\int_{x_{j-1}}^{x_j} |\eta(x)(t)|^2\,dt = \begin{cases} 0, & \text{with probability } p_{2j-1}, \\ \displaystyle\int_{x_{j-1}}^{x_j} \frac{\tau_{2j-1}^2}{(1 - p_{2j-1})^2}\,dt, & \text{with probability } 1 - p_{2j-1}. \end{cases}$$

Therefore,

$$\mathbb{M}\left(\int_{x_{j-1}}^{x_j} |\eta(x)(t)|^2\,dt\right) = \int_{x_{j-1}}^{x_j} \frac{\tau_{2j-1}^2}{1 - p_{2j-1}}\,dt.$$

At the same time

$$\int_{x_{j-1}}^{x_j} \mathbb{M}|\eta(x)(t)|^2\,dt = \int_{x_{j-1}}^{x_j} \frac{\tau_{2j-1}^2}{1 - p_{2j-1}}\,dt.$$

The proof for the interval $(-x_j, -x_{j-1})$ is completely similar. Consequently, $\eta(x)(\cdot)$ satisfies conditions (1) and (2). Thus, $\eta(x)(\cdot) \in Y_\delta(x(\cdot))$ for all $x(\cdot) \in B$.

Let φ be an arbitrary recovery method. Taking into account that the set B

is finite (with 2^{2N} elements), we have

$$
\begin{aligned}
e^2(\Lambda, W, \delta, \varphi) &\geq \sup_{x(\cdot)\in B} \mathbb{M}\|\Lambda x(\cdot) - \varphi(\eta(x)(\cdot))(\cdot)\|^2_{L_2(\mathbb{R})} \\
&= \sup_{x(\cdot)\in B}\left(\sum_{j=1}^{2N+1}(p_j - p_{j-1})\left\|\Lambda x(\cdot) - \varphi\Big(\sum_{k=1}^{j-1}\frac{s_k(x)\tau_k}{1-p_k}e_k(\cdot)\Big)(\cdot)\right\|^2_{L_2(\mathbb{R})}\right) \\
&\geq \frac{1}{2^{2N}}\sum_{x(\cdot)\in B}\left(\sum_{j=1}^{2N+1}(p_j - p_{j-1})\left\|\Lambda x(\cdot) - \varphi\Big(\sum_{k=1}^{j-1}\frac{s_k(x)\tau_k}{1-p_k}e_k(\cdot)\Big)(\cdot)\right\|^2_{L_2(\mathbb{R})}\right) \\
&= \frac{1}{2^{2N}}\sum_{j=1}^{2N+1}(p_j - p_{j-1})\sum_{x(\cdot)\in B}\left\|\Lambda x(\cdot) - \varphi\Big(\sum_{k=1}^{j-1}\frac{s_k(x)\tau_k}{1-p_k}e_k(\cdot)\Big)(\cdot)\right\|^2_{L_2(\mathbb{R})}; \quad (6)
\end{aligned}
$$

here $p_0 = 0$ and $p_{2N+1} = 1$. Set

$$
B_{s_1,\ldots,s_{j-1}} = \{\, x(\cdot) \in B : s_1(x) = s_1, \ldots, s_{j-1}(x) = s_{j-1}\,\},
$$

$j = 1, \ldots, 2N+1$ (for $j = 1$ this set coincides with B). Then

$$
\begin{aligned}
\frac{p_j - p_{j-1}}{2^{2N}}\sum_{x(\cdot)\in B}\left\|\Lambda x(\cdot) - \varphi\Big(\sum_{k=1}^{j-1}\frac{s_k(x)\tau_k}{1-p_k}e_k(\cdot)\Big)(\cdot)\right\|^2_{L_2(\mathbb{R})} &= \frac{p_j - p_{j-1}}{2^{2N}} \\
\times \sum_{s_1,\ldots,s_{j-1}}\ \sum_{x\in B_{s_1,\ldots,s_{j-1}}}\left\|\Lambda x(\cdot) - \varphi\Big(\sum_{k=1}^{j-1}\frac{s_k\tau_k}{1-p_k}e_k(\cdot)\Big)(\cdot)\right\|^2_{L_2(\mathbb{R})}&.
\end{aligned}
$$

If $x(\cdot) \in B_{s_1,\ldots,s_{j-1}}$, then

$$
x(\cdot) = \sum_{k=1}^{j-1} s_k\tau_k e_k(\cdot) + z(x)(\cdot), \quad z(x)(\cdot) = \sum_{k=j}^{2N} s_k(x)\tau_k e_k(\cdot).
$$

Moreover, with every element

$$
\sum_{k=1}^{j-1} s_k\tau_k e_k(\cdot) + z(x)(\cdot) \in B_{s_1,\ldots,s_{j-1}}
$$

the set $B_{s_1,\ldots,s_{j-1}}$ contains the element

$$
\sum_{k=1}^{j-1} s_k\tau_k e_k(\cdot) - z(x)(\cdot).
$$

Thus,

$$
\begin{aligned}
&\frac{p_j-p_{j-1}}{2^{2N}}\sum_{s_1,\dots,s_{j-1}}\ \sum_{x\in B_{s_1,\dots,s_{j-1}}}\left\|\Lambda x(\cdot)-\varphi\Bigg(\sum_{k=1}^{j-1}\frac{s_k\tau_k}{1-p_k}e_k(\cdot)\Bigg)(\cdot)\right\|^2_{L_2(\mathbb{R})}\\
&\quad=\frac{p_j-p_{j-1}}{2^{2N}}\sum_{s_1,\dots,s_{j-1}}\ \sum_{x\in B_{s_1,\dots,s_{j-1}}}\Bigg\|\Lambda\Bigg(\sum_{k=1}^{j-1}s_k\tau_ke_k(\cdot)+z(x)(\cdot)\Bigg)\\
&\qquad\qquad-\varphi\Bigg(\sum_{k=1}^{j-1}\frac{s_k\tau_k}{1-p_k}e_k(\cdot)\Bigg)(\cdot)\Bigg\|^2_{L_2(\mathbb{R})}\\
&\quad=\frac{p_j-p_{j-1}}{2^{2N}}\sum_{s_1,\dots,s_{j-1}}\ \sum_{x\in B_{s_1,\dots,s_{j-1}}}\Bigg\|\Lambda\Bigg(\sum_{k=1}^{j-1}s_k\tau_ke_k(\cdot)\Bigg)+\Lambda z(x)(\cdot)\\
&\qquad\qquad-\varphi\Bigg(\sum_{k=1}^{j-1}\frac{s_k\tau_k}{1-p_k}e_k(\cdot)\Bigg)(\cdot)\Bigg\|^2_{L_2(\mathbb{R})}\\
&\quad=\frac{p_j-p_{j-1}}{2^{2N+1}}\sum_{s_1,\dots,s_{j-1}}\ \sum_{x\in B_{s_1,\dots,s_{j-1}}}\Bigg(\Bigg\|\Lambda\Bigg(\sum_{k=1}^{j-1}s_k\tau_ke_k(\cdot)\Bigg)+\Lambda z(x)(\cdot)\\
&\quad-\varphi\Bigg(\sum_{k=1}^{j-1}\frac{s_k\tau_k}{1-p_k}e_k(\cdot)\Bigg)(\cdot)\Bigg\|^2_{L_2(\mathbb{R})}+\Bigg\|\Lambda\Bigg(\sum_{k=1}^{j-1}s_k\tau_ke_k(\cdot)\Bigg)-\Lambda z(x)(\cdot)\\
&\qquad\qquad-\varphi\Bigg(\sum_{k=1}^{j-1}\frac{s_k\tau_k}{1-p_k}e_k(\cdot)\Bigg)(\cdot)\Bigg\|^2_{L_2(\mathbb{R})}\Bigg)\\
&\qquad\geq\frac{p_j-p_{j-1}}{2^{2N}}\sum_{s_1,\dots,s_{j-1}}\ \sum_{x\in B_{s_1,\dots,s_{j-1}}}\|\Lambda z(x)(\cdot)\|^2_{L_2(\mathbb{R})}\\
&\qquad\qquad=\frac{p_j-p_{j-1}}{2^{2N}}\sum_{x\in B}\|\Lambda z(x)(\cdot)\|^2_{L_2(\mathbb{R})}=(p_j-p_{j-1})\sum_{k=j}^{2N}\mu_k\tau_k^2,
\end{aligned}
$$

where

$$
\mu_{2j-1}=\int_{x_{j-1}}^{x_j}|\mu(t)|^2\,dt,\quad \mu_{2j}=\int_{-x_j}^{-x_{j-1}}|\mu(t)|^2\,dt,\quad j=1,\dots,N.
$$

Substituting this estimate into (6), we get

$$
\begin{aligned}
e^2(\Lambda,W,\delta,\varphi)&\geq\sum_{j=1}^{2N+1}(p_j-p_{j-1})\sum_{k=j}^{2N}\mu_k\tau_k^2\\
&=\sum_{j=1}^{2N}\Bigg(p_j\sum_{k=j}^{2N}\mu_k\tau_k^2-p_j\sum_{k=j+1}^{2N}\mu_k\tau_k^2\Bigg)=\sum_{j=1}^{2N}p_j\mu_j\tau_j^2=\sum_{j=1}^{2N}\frac{\delta^2}{\delta^2+\tau_j^2}\mu_j\tau_j^2.
\end{aligned}
$$

Since the method φ was chosen arbitrarily, we have

$$E^2(\Lambda, W, \delta) \geq \sup_{\substack{\tau_1 \geq \ldots \geq \tau_{2N} > 0 \\ x_\tau(\cdot) \in W}} \sum_{j=1}^{2N} \frac{\delta^2}{\delta^2 + \tau_j^2} \mu_j \tau_j^2. \tag{7}$$

The condition $x_\tau(\cdot) \in W$ means that

$$\int_{\mathbb{R}} \nu(t) |x_\tau(t)|^2 \, dt = \sum_{j=1}^{2N} \nu_j \tau_j^2 \leq 1,$$

where

$$\nu_{2j-1} = \int_{x_{j-1}}^{x_j} \nu(t) \, dt, \quad \nu_{2j} = \int_{-x_j}^{-x_{j-1}} \nu(t) \, dt, \quad j = 1, \ldots, N.$$

Hence,

$$E^2(\Lambda, W, \delta) \geq \sup_{\substack{\tau_1 \geq \ldots \geq \tau_{2N} > 0 \\ \sum_{j=1}^{2N} \nu_j \tau_j^2 \leq 1}} \sum_{j=1}^{2N} \frac{\delta^2}{\delta^2 + \tau_j^2} \mu_j \tau_j^2.$$

Let $\tau = (\tau_1, \ldots, \tau_k, 0, \ldots, 0)$, $1 \leq k < 2N$, $\tau_1 \geq \ldots \geq \tau_k > 0$, and

$$\sum_{j=1}^{k} \nu_j \tau_j^2 \leq 1.$$

For sufficiently small $\varepsilon > 0$ we put $\tau_\varepsilon = (\tau_1(\varepsilon), \ldots, \tau_{2N}(\varepsilon))$ where

$$\tau_j(\varepsilon) = \begin{cases} \sqrt{\tau_j^2 - \varepsilon}, & 1 \leq j \leq k, \\ C\sqrt{\varepsilon}, & k+1 \leq j \leq 2N, \end{cases}$$

and

$$C = \left(\frac{\sum_{j=1}^{k} \nu_j}{\sum_{j=k+1}^{2N} \nu_j} \right)^{1/2}.$$

Then

$$\sum_{j=1}^{2N} \nu_j \tau_j^2(\varepsilon) = \sum_{j=1}^{k} \nu_j \tau_j^2 - \varepsilon \sum_{j=1}^{k} \nu_j + C^2 \varepsilon \sum_{j=k+1}^{2N} \nu_j = \sum_{j=1}^{k} \nu_j \tau_j^2 \leq 1.$$

For $\varepsilon < \tau_k^2/(1 + C^2)$ we have

$$\sqrt{\tau_k^2 - \varepsilon} > C\sqrt{\varepsilon}.$$

Consequently, for such ε

$$\tau_1(\varepsilon) \geq \ldots \geq \tau_{2N}(\varepsilon) > 0.$$

It follows from (7) that

$$E^2(\Lambda, W, \delta) \geq \sum_{j=1}^{2N} \frac{\delta^2}{\delta^2 + \tau_j^2(\varepsilon)} \mu_j \tau_j^2(\varepsilon).$$

Passing to the limit as $\varepsilon \to 0$, we obtain

$$E^2(\Lambda, W, \delta) \geq \sum_{j=1}^{k} \frac{\delta^2}{\delta^2 + \tau_j^2} \mu_j \tau_j^2.$$

Thus,

$$E^2(\Lambda, W, \delta) \geq \sup_{\substack{\tau_1 \geq \ldots \geq \tau_{2N} \geq 0 \\ \sum_{j=1}^{2N} \nu_j \tau_j^2 \leq 1}} \sum_{j=1}^{2N} \frac{\delta^2}{\delta^2 + \tau_j^2} \mu_j \tau_j^2. \tag{8}$$

Let the piecewise continuous function $x(\cdot)$ be such that $|x(\cdot)|$ is an even function monotonically decreasing on $\mathbb{R}_+$ and

$$\int_{\mathbb{R}} \nu(t)|x(t)|^2 \, dt < 1. \tag{9}$$

Consider the integral

$$I = \int_{\mathbb{R}} \frac{\delta^2}{\delta^2 + |x(t)|^2} |\mu(t)|^2 |x(t)|^2 \, dt.$$

Let us fix $\varepsilon > 0$ and find $A > 0$, $0 < \widetilde{a} < A$ such that

$$I_1 = \int_{[-A, -\widetilde{a}] \cup [\widetilde{a}, A]} \frac{\delta^2}{\delta^2 + |x(t)|^2} |\mu(t)|^2 |x(t)|^2 \, dt > I - \varepsilon. \tag{10}$$

It is obvious that

$$I_2 = \int_{[-A, -\widetilde{a}] \cup [\widetilde{a}, A]} \nu(t)|x(t)|^2 \, dt < 1.$$

We will approximate these integrals by integral sums over partitions of T_N, representing segments $[x_{j-1}, x_j]$, $[-x_j, -x_{j-1}]$, $j = 1, \ldots, N$, and points

$$t_{2j-1} = \widetilde{a} + (2j-1)\frac{A - \widetilde{a}}{2N}, \quad t_{2j} = -t_{2j-1}, \quad j = 1, \ldots, N.$$

For any $\varepsilon_1 > 0$, there is such a N_1 that for all $N > N_1$

$$\frac{1}{N} \sum_{j=1}^{2N} \frac{\delta^2}{\delta^2 + \tau_j^2} |\mu(t_j)|^2 \tau_j^2 > I_1 - \varepsilon_1, \tag{11}$$

where $\tau_j = |x(t_j)|$. Moreover, there is such a N_2 that for all $N > N_2$ for some $\omega > 0$ the inequality

$$\frac{1}{N}\sum_{j=1}^{2N} \nu(t_j)\tau_j^2 < 1 - \omega$$

holds.

By the mean value theorem for integrals, there are ξ_j such that $\xi_{2j-1} \in [x_{j-1}, x_j]$, $\xi_{2j} \in [-x_j, -x_{j-1}]$, $j = 1, \ldots, N$, and $\mu_j = |\mu(\xi_j)|^2/N$. Using the same arguments we obtain that there are η_j such that $\eta_{2j-1} \in [x_{j-1}, x_j]$, $\eta_{2j} \in [-x_j, -x_{j-1}]$, $j = 1, \ldots, N$, and $\nu_j = \nu(\eta_j)/N$. Due to the uniform continuity of the functions $|\mu(\cdot)|$ and $\nu(\cdot)$ on the segment $[-A, A]$ for any $\varepsilon_2 > 0$, there is N_3 such that for all $s_1, s_2 \in [-A, A]$, $|s_1 - s_2| < 1/N_3$, inequalities

$$||\mu(s_1)|^2 - |\mu(s_2)|^2| < \varepsilon_2, \quad |\nu(s_1) - \nu(s_2)| < \varepsilon_2 \tag{12}$$

hold.

Put

$$M = \operatorname*{vraisup}_{t\in[\widetilde{a},A]} |x(t)|^2$$

(due to the monotonous decrease of $|x(\cdot)|$ on $\mathbb{R}_+$, $M < \infty$). Choose $\varepsilon_2 < \omega/(2M)$. Let $N > \max\{N_1, N_2, N_3\}$. Then $\nu_j = \nu(\eta_j)/N < \nu(t_j)/N + \varepsilon_2/N$. Consequently,

$$\sum_{j=1}^{2N} \nu_j\tau_j^2 \le \frac{1}{N}\sum_{j=1}^{2N}(\nu(t_j) + \varepsilon_2)\tau_j^2 \le \frac{1}{N}\sum_{j=1}^{2N} \nu(t_j)\tau_j^2 + 2M\varepsilon_2 < 1. \tag{13}$$

It follows from (12) that $\mu_j = |\mu(\xi_j)|^2/N > |\mu(t_j)|^2/N - \varepsilon_2/N$. Therefore,

$$\sum_{j=1}^{2N} \frac{\delta^2}{\delta^2 + \tau_j^2}\mu_j\tau_j^2 \ge \frac{1}{N}\sum_{j=1}^{2N} \frac{\delta^2}{\delta^2 + \tau_j^2}(|\mu(t_j)|^2 - \varepsilon_2)\tau_j^2$$
$$\ge \frac{1}{N}\sum_{j=1}^{2N} \frac{\delta^2}{\delta^2 + \tau_j^2}|\mu(t_j)|^2\tau_j^2 - 2M\varepsilon_2.$$

Taking into account (13) and (11), it follows from (8) that

$$E^2(\Lambda, W, \delta) \ge \sum_{j=1}^{2N} \frac{\delta^2}{\delta^2 + \tau_j^2}\mu_j\tau_j^2 \ge \frac{1}{N}\sum_{j=1}^{2N} \frac{\delta^2}{\delta^2 + \tau_j^2}|\mu(t_j)|^2\tau_j^2 - 2M\varepsilon_2$$
$$\ge I_1 - \varepsilon_1 - 2M\varepsilon_2.$$

Due to the fact that ε_1 and ε_2 can be chosen arbitrarily small, we get

$$E^2(\Lambda, W, \delta) \ge I_1 \ge I - \varepsilon.$$

Since ε can be chosen arbitrarily small, we obtain

$$E^2(\Lambda, W, \delta) \geq \sup_{\substack{x(\cdot)\in W_0 \\ \int_{\mathbb{R}} \nu(t)|x(t)|^2\,dt<1}} \int_{\mathbb{R}} \frac{\delta^2}{\delta^2 + |x(t)|^2} |\mu(t)|^2 |x(t)|^2\,dt, \qquad (14)$$

where W_0 is the set of piecewise continuous functions $x(\cdot)$ such that $|x(\cdot)|$ is an even function monotonically decreasing on $\mathbb{R}_+$.

We show that the strict inequality on the right side of (14) can be replaced by a non-strict one. Let $x(\cdot) \in W_0$ and

$$\int_{\mathbb{R}} \nu(t)|x(t)|^2\,dt = 1.$$

Consider the function $y(\cdot) = (1+\varepsilon)^{-1/2}x(\cdot)$, $\varepsilon > 0$. Then it follows from (14) that

$$E^2(\Lambda, W, \delta) \geq \frac{1}{1+\varepsilon} \int_{\mathbb{R}} \frac{\delta^2}{\delta^2 + \dfrac{|x(t)|^2}{1+\varepsilon}} |\mu(t)|^2 |x(t)|^2\,dt.$$

Since

$$\frac{1}{\delta^2 + \dfrac{|x(t)|^2}{1+\varepsilon}} \geq \frac{1}{\delta^2 + |x(t)|^2},$$

we have

$$E^2(\Lambda, W, \delta) \geq \frac{1}{1+\varepsilon} \int_{\mathbb{R}} \frac{\delta^2}{\delta^2 + |x(t)|^2} |\mu(t)|^2 |x(t)|^2\,dt.$$

Passing ε to zero, we get

$$E^2(\Lambda, W, \delta) \geq \int_{\mathbb{R}} \frac{\delta^2}{\delta^2 + |x(t)|^2} |\mu(t)|^2 |x(t)|^2\,dt.$$

Thus,

$$E^2(\Lambda, W, \delta) \geq \sup_{\substack{x(\cdot)\in W_0 \\ \int_{\mathbb{R}} \nu(t)|x(t)|^2\,dt\leq 1}} \int_{\mathbb{R}} \frac{\delta^2}{\delta^2 + |x(t)|^2} |\mu(t)|^2 |x(t)|^2\,dt. \qquad (15)$$

2. The upper bound. Let us find the error of the methods having the form

$$\varphi(y_\xi(\cdot))(\cdot) = \alpha(\cdot)\mu(\cdot)y_\xi(\cdot).$$

Put $z_\xi(\cdot) = y_\xi(\cdot) - x(\cdot)$. Then $\mathbb{M}z_\xi(\cdot) = 0$, $\operatorname{Var} z_\xi(\cdot) \leq \delta^2$. Using the well-known bias-variance decomposition we have

$$e^2(\Lambda, W, \delta, \varphi) = \sup_{\substack{x(\cdot)\in W \\ y_\xi(\cdot)\in Y_\delta(x(\cdot))}} \mathbb{M}\left(\|\Lambda x(\cdot) - \varphi(y_\xi(\cdot))(\cdot)\|^2_{L_2(\mathbb{R})}\right)$$

$$= \sup_{\substack{x(\cdot)\in W \\ y_\xi(\cdot)\in Y_\delta(x(\cdot))}} \mathbb{M}\left(\|\Lambda x(\cdot) - \varphi(x(\cdot))(\cdot) - \varphi(z_\xi(\cdot))(\cdot)\|^2_{L_2(\mathbb{R})}\right)$$

$$= \sup_{\substack{x(\cdot)\in W \\ y_\xi(\cdot)\in Y_\delta(x(\cdot))}} \Big(\|\Lambda x(\cdot) - \varphi(x(\cdot))(\cdot)\|^2_{L_2(\mathbb{R})} + \mathbb{M}(\|\varphi(z_\xi(\cdot))(\cdot)\|^2_{L_2(\mathbb{R})})$$

$$-2\mathbb{M}\,\mathrm{Re}(\varphi(z_\xi(\cdot))(\cdot), \overline{\Lambda x(\cdot) - \varphi(x(\cdot))(\cdot)})\Big);$$

here $(\cdot,\cdot)$ is the standard scalar product in $L_2(\mathbb{R})$. It follows from the form of φ that

$$\mathbb{M}\,\mathrm{Re}(\varphi(z_\xi(\cdot))(\cdot), \overline{\Lambda x(\cdot) - \varphi(x(\cdot))(\cdot)})$$

$$= \mathrm{Re}\,\mathbb{M}\left(\alpha(\cdot)\mu(\cdot)z_\xi(\cdot), \overline{\Lambda x(\cdot) - \varphi(x(\cdot))(\cdot)}\right)$$

$$= \mathrm{Re}\left(\alpha(\cdot)\mu(\cdot)\mathbb{M}z_\xi(\cdot), \overline{\Lambda x(\cdot) - \varphi(x(\cdot))(\cdot)}\right) = 0.$$

Due to the fact that $\mathbb{M}(|z_\xi(\cdot)|^2) = \mathrm{Var}\, y_\xi(\cdot)$, we have

$$e^2(\Lambda, W, \delta, \varphi) = \sup_{\substack{x(\cdot)\in W \\ y_\xi(\cdot)\in Y_\delta(x(\cdot))}} \left(\int_{\mathbb{R}} |\mu(t)|^2|1-\alpha(t)|^2|x(t)|^2\,dt\right.$$

$$\left. + \int_{\mathbb{R}} |\mu(t)|^2|\alpha(t)|^2\,\mathrm{Var}\,y_\xi(t)\,dt\right) = \sup_{x(\cdot)\in W} \int_{\mathbb{R}} |\mu(t)|^2|1-\alpha(t)|^2|x(t)|^2\,dt$$

$$+ \delta^2\int_{\mathbb{R}} |\mu(t)|^2|\alpha(t)|^2\,dt$$

Since

$$\int_{\mathbb{R}} |\mu(t)|^2|1-\alpha(t)|^2|x(t)|^2\,dt = \int_{\mathbb{R}} \frac{|\mu(t)|^2}{\nu(t)}|1-\alpha(t)|^2\nu(t)|x(t)|^2\,dt$$

$$\le \operatorname*{vraisup}_{t\in\mathbb{R}} \left(\frac{|\mu(t)|^2}{\nu(t)}|1-\alpha(t)|^2\right),$$

we obtain

$$e^2(\Lambda, W, \delta, \varphi) \le \operatorname*{vraisup}_{t\in\mathbb{R}} \left(\frac{|\mu(t)|^2}{\nu(t)}|1-\alpha(t)|^2\right) + \delta^2\int_{\mathbb{R}} |\mu(t)|^2|\alpha(t)|^2\,dt.$$

Put

$$\alpha(t) = \left(1 - \frac{\sqrt{\nu(t)}}{|\mu(t)|}\frac{|\mu(t_\delta)|}{\sqrt{\nu(t_\delta)}}\right)_+.$$

Due to the monotonous decreasing of the function $|\mu(\cdot)|/\sqrt{\nu(\cdot)}$ we get

$$\operatorname*{vraisup}_{t\in\mathbb{R}}\left(\frac{|\mu(t)|^2}{\nu(t)}|1-\alpha(t)|^2\right)=\frac{|\mu(t_\delta)|^2}{\nu(t_\delta)}.$$

Consequently,

$$\begin{aligned}
e^2(\Lambda,W,\delta,\varphi)\le{}&\frac{|\mu(t_\delta)|^2}{\nu(t_\delta)}\\
&+\delta^2\int_{|t|\le t_\delta}|\mu(t)|^2\left(1-\frac{\sqrt{\nu(t)}}{|\mu(t)|}\frac{|\mu(t_\delta)|}{\sqrt{\nu(t_\delta)}}\right)^2dt=\frac{|\mu(t_\delta)|^2}{\nu(t_\delta)}\\
&+\delta^2\int_{|t|\le t_\delta}|\mu(t)|^2\left(\left(1-\frac{\sqrt{\nu(t)}}{|\mu(t)|}\frac{|\mu(t_\delta)|}{\sqrt{\nu(t_\delta)}}\right)-\frac{\sqrt{\nu(t)}}{|\mu(t)|}\frac{|\mu(t_\delta)|}{\sqrt{\nu(t_\delta)}}\right.\\
&\left.+\frac{\nu(t)}{|\mu(t)|^2}\frac{|\mu(t_\delta)|^2}{\nu(t_\delta)}\right)dt=\delta^2\int_{|t|\le t_\delta}|\mu(t)|^2\left(1-\frac{\sqrt{\nu(t)}}{|\mu(t)|}\frac{|\mu(t_\delta)|}{\sqrt{\nu(t_\delta)}}\right)dt\\
&+\frac{|\mu(t_\delta)|}{\sqrt{\nu(t_\delta)}}\left(\frac{|\mu(t_\delta)|}{\sqrt{\nu(t_\delta)}}\left(1+\delta^2\int_{|t|\le t_\delta}\nu(t)\,dt\right)-\delta^2\int_{|t|\le t_\delta}|\mu(t)|\sqrt{\nu(t)}\,dt\right)\\
&=\delta^2\int_{|t|\le t_\delta}|\mu(t)|^2\left(1-\frac{\sqrt{\nu(t)}}{|\mu(t)|}\frac{|\mu(t_\delta)|}{\sqrt{\nu(t_\delta)}}\right)dt+\frac{|\mu(t_\delta)|^2}{\nu(t_\delta)}(1-\delta^2f(t_\delta).
\end{aligned}$$

It follows from the definition of t_δ that $f(t_\delta)=\delta^{-2}$. Thus,

$$e^2(\Lambda,W,\delta,\varphi)\le\delta^2\int_{|t|\le t_\delta}|\mu(t)|^2\left(1-\frac{\sqrt{\nu(t)}}{|\mu(t)|}\frac{|\mu(t_\delta)|}{\sqrt{\nu(t_\delta)}}\right)dt. \tag{16}$$

Consider the function

$$\widehat{x}(t)=\delta\left(\left(\frac{\sqrt{\nu(t_\delta)}}{|\mu(t_\delta)|}\frac{|\mu(t)|}{\sqrt{\nu(t)}}-1\right)_+\right)^{1/2}.$$

It is obvious that $\widehat{x}(\cdot)\in W_0$. Moreover,

$$\int_{\mathbb{R}}\nu(t)|\widehat{x}(t)|^2\,dt=\delta^2\int_{|t|\le t_\delta}\nu(t)\left(\frac{\sqrt{\nu(t_\delta)}}{|\mu(t_\delta)|}\frac{|\mu(t)|}{\sqrt{\nu(t)}}-1\right)dt=1.$$

Taking into account (16), it follows from (15) that

$$\begin{aligned}
E^2(\Lambda,W,\delta)\ge{}&\int_{\mathbb{R}}\frac{\delta^2}{\delta^2+|\widehat{x}(t)|^2}|\mu(t)|^2|\widehat{x}(t)|^2\,dt\\
&=\delta^2\int_{|t|\le t_\delta}|\mu(t)|^2\left(1-\frac{\sqrt{\nu(t)}}{|\mu(t)|}\frac{|\mu(t_\delta)|}{\sqrt{\nu(t_\delta)}}\right)dt\\
&\ge e^2(\Lambda,W,\delta,\varphi)\ge E^2(\Lambda,W,\delta).
\end{aligned}$$

This implies (5) and the optimality of the method φ. □

Now we consider some examples of the application of Theorem 1.

3. Recovery of functions and their derivatives from the Fourier transform given with random error

Denote by $\mathcal{W}_2^r(\mathbb{R})$ the set of functions $x(\cdot) \in L_2(\mathbb{R})$ for which $x^{(r-1)}(\cdot)$ is locally absolutely continuous and $x^{(r)}(\cdot) \in L_2(\mathbb{R})$. Put

$$W_2^r(\mathbb{R}) = \{\, x(\cdot) \in \mathcal{W}_2^r(\mathbb{R}) : \|x^{(r)}(\cdot)\|_{L_2(\mathbb{R})} \le 1 \,\}.$$

Suppose that the Fourier transform $Fx(\cdot)$ of the function $x(\cdot) \in W_2^r(\mathbb{R})$ is given with a random error. We assume that instead of the function $Fx(\cdot)$ we know a random function $y_\xi(\cdot) \in L_2(\mathbb{R})$ such that $\mathbb{M} y_\xi(\cdot) = Fx(\cdot)$ and $\operatorname{Var} y_\xi(\cdot) \le \delta^2$ almost everywhere. Using this information, it is required to recover the function $D^k x(\cdot) = x^{(k)}(\cdot)$, $0 \le k < r$, in the $L_2(\mathbb{R})$-metric.

The exact setting of the problem is as follows. For every $x(\cdot) \in W_2^r(\mathbb{R})$ we consider the set of random functions

$$Y_\delta^F(x(\cdot)) = \{\, y_\xi(\cdot) \in L_2(\mathbb{R}) : \mathbb{M} y_\xi(\cdot) = Fx(\cdot),\ \operatorname{Var} y_\xi(\cdot) \le \delta^2 \text{ a.e.} \,\}.$$

This set requires additional conditions given in the general setting (the validity of the equalities (1) and (2)). Next, we define the error of the recovery method $\varphi\colon L_2(\mathbb{R}) \to L_2(\mathbb{R})$ as follows

$$e(D^k, W_2^r(\mathbb{R}), \delta, \varphi) = \left(\sup_{\substack{x(\cdot) \in W_2^r(\mathbb{R}) \\ y_\xi(\cdot) \in Y_\delta^F(x(\cdot))}} \mathbb{M} \left(\|x^{(k)}(\cdot) - \varphi(y_\xi(\cdot))(\cdot)\|_{L_2(\mathbb{R})}^2 \right) \right)^{1/2}.$$

The problem is to find the error of optimal recovery

$$E(D^k, W_2^r(\mathbb{R}), \delta) = \inf_{\varphi\colon L_2(\mathbb{R}) \to L_2(\mathbb{R})} e(D^k, W_2^r(\mathbb{R}), \delta, \varphi)$$

and a method on which this infimum is attained.

It follows from the Parseval equality that

$$\|x^{(r)}(\cdot)\|_{L_2(\mathbb{R})}^2 = \frac{1}{2\pi} \int_{\mathbb{R}} |t|^{2r} |Fx(t)|^2 \, dt,$$

$$\|x^{(k)}(\cdot) - \varphi(y_\xi(\cdot))(\cdot)\|_{L_2(\mathbb{R})}^2 = \frac{1}{2\pi} \int_{\mathbb{R}} |(it)^k F(t) - F\varphi(y_\xi(\cdot))(t)|^2 \, dt.$$

Thus, the problem is reduced to problem (3) with $\nu(t) = t^{2r}$ and $\mu(t) = (it)^k$. The function $f(\cdot)$ which was defined by (4), has the form

$$f(s) = \int_{|t| \le s} \left(\frac{s^{r-k}}{|t|^{r-k}} - 1 \right) t^{2r} \, dt = 2 s^{2r+1} \frac{r-k}{(2r+1)(r+k+1)}.$$

The equation $f(s) = \delta^{-2}$ has the unique solution

$$t_\delta = \left(\frac{(2r+1)(r+k+1)}{2\delta^2 (r-k)} \right)^{\frac{1}{2r+1}}.$$

It follows from Theorem 1

Theorem 2. *For all $\delta > 0$ and $0 \le k < r$*

$$E(D^k, W_2^r(\mathbb{R}), \delta) = \frac{(2r+1)^{\frac{2k+1}{2(2r+1)}}}{\sqrt{2k+1}} \left(2\delta^2 \frac{r-k}{r+k+1}\right)^{\frac{r-k}{2(2r+1)}}.$$

Moreover, the method

$$\varphi(y_\xi(\cdot))(t) = F^{-1}\left((it)^k \alpha(t) y_\xi(t)\right)(t),$$

where

$$\alpha(t) = \left(1 - t^{r-k}\left(\frac{2\delta^2(r-k)}{(2r+1)(r+k+1)}\right)^{\frac{r-k}{2r+1}}\right)_+,$$

is optimal.

Note that the optimal recovery method does not use all the information about random functions $y_\xi(\cdot)$, but only the information contained in the segment $[-t_\delta, t_\delta]$. Moreover, the more accurate the measurements (the smaller the variance δ^2), the larger this segment becomes.

The deterministic case of this problem was considered in [5] (see also [7]).

4. Recovery of the solution of the heat equation

The temperature distribution in an infinite rod is described by the equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial t^2},$$

where $u(\cdot,\cdot)$ is the function on $[0,\infty) \times \mathbb{R}$ with a given initial temperature distribution

$$u(0,\cdot) = u_0(\cdot).$$

Consider the problem of recovering the temperature distribution at an instant of time T from information about the Fourier transform of the initial temperature distribution $u_0(\cdot)$, given with a random error. We assume that the functions $u_0(\cdot)$, given the initial temperature distribution, belong to the class $W_2^r(\mathbb{R})$. We define the error of the recovery method $\varphi\colon L_2(\mathbb{R}) \to L_2(\mathbb{R})$ as follows

$$e(T, W_2^r(\mathbb{R}), \delta, \varphi) = \left(\sup_{\substack{u_0(\cdot)\in W_2^r(\mathbb{R}) \\ y_\xi(\cdot)\in Y_\delta^F(u_0(\cdot))}} \mathbb{M}\left(\|u(T,\cdot) - \varphi(y_\xi(\cdot))(\cdot)\|^2_{L_2(\mathbb{R})}\right)\right)^{1/2}.$$

The problem is to find the error of optimal recovery

$$E(T, W_2^r(\mathbb{R}), \delta) = \inf_{\varphi\colon L_2(\mathbb{R})\to L_2(\mathbb{R})} e(T, W_2^r(\mathbb{R}), \delta, \varphi)$$

and a method on which this infimum is attained.

It is well known (see, for example, [3]) that for all $t \geq 0$ the equality

$$F(u(t,\cdot))(\lambda) = e^{-\lambda^2 t} F u_0(\lambda)$$

holds. It follows from the Parseval equality that

$$\|u(T,\cdot) - \varphi(y_\xi(\cdot))(\cdot)\|^2_{L_2(\mathbb{R})} = \frac{1}{2\pi} \int_{\mathbb{R}} |e^{-\lambda^2 T} F u_0(\lambda) - F\varphi(y_\xi(\cdot))(\lambda)|^2 \, d\lambda.$$

Thus, the problem is reduced to problem (3) with $\nu(t) = t^{2r}$ and $\mu(t) = e^{-t^2 T}$. The function $f(\cdot)$ which was defined by (4), has the form

$$f(s) = \int_{|t| \leq s} \left(\frac{s^r e^{s^2 T}}{t^r e^{t^2 T}} - 1 \right) t^{2r} \, dt = 2 s^r e^{s^2 T} \int_0^s t^r e^{-t^2 T} \, dt - \frac{s^{2r+1}}{2r+1}.$$

It is easy to verify that $f(s) \to +\infty$ as $s \to +\infty$ (the monotonous increase of $f(\cdot)$ was noted in the general case). Therefore, the equation $f(s) = \delta^{-2}$ has a unique solution, which we denote by t_δ.

From Theorem 1 we obtain the following result:

Theorem 3. *Put*

$$\alpha(t) = \left(1 - \frac{t^r e^{t^2 T}}{t_\delta^r e^{t_\delta^2 T}} \right)_+ .$$

Then the equality

$$E(T, W_2^r(\mathbb{R}), \delta) = \delta \left(\int_{|t| \leq t_\delta} e^{-2t^2 T} \alpha(t) \, dt \right)^{1/2}$$

holds. Moreover, the method

$$\varphi(y_\xi(\cdot))(t) = F^{-1} \left(e^{-t^2 T} \alpha(t) y_\xi(t) \right)(t)$$

is optimal.

The deterministic case of this problem was considered in [6].

[1] D. L. Donoho, Statistical estimation and optimal recovery, Ann. Statist. 22 (1) (1994) 238–270.

[2] D. L. Donoho, C. Liu, B. MacGibbon, Minimax risk over hyperrectangles, and implications, Ann. Statist. 18 (3) (1990) 1416–1437.

[3] A. N. Kolmogorov, S. V. Fomin, Elements of theory of functions and functional analysis, Nauka, Moscow, 1972 (in Russion).

[4] K. Yu. Krivosheev, On optimal recovery of values of linear operators from information known with a stochastic error, Sb. Math. 212 (11) (2021) 1588–1607.

[5] G. G. Magaril-Il'yaev, K. Yu. Osipenko, Optimal recovery of functions and their derivatives from inaccurate information about a spectrum and inequalities for derivatives, Funct. Anal and Its Appl. 37 (3) (2003) 203–214.

[6] G. G. Magaril-Il'yaev, K. Yu. Osipenko, Optimal recovery of the solution of the heat equation from inaccurate data, Sb. Math. 200 (5) (2009) 665–682.

[7] G. G. Magaril-Il'yaev, K. Yu. Osipenko, On optimal harmonic synthesis from inaccurate spectral data, Funct. Anal and Its Appl. 44 (3) (2010) 223–225.

[8] G. G. Magaril-Il'yaev, V. M. Tichomirov, Convex Analysis: Theory and Applications, Translations of Math. Monographs, vol. 222, AMS, Providence, RI, 2003.

[9] C. A. Micchelli, T. J. Rivlin, A survey of optimal recovery, in Optimal Estimation in Approximation Theory, C. A. Micchelli and T. J. Rivlin, eds., Plenum Press, New York, 1977, pp. 1–54.

[10] K. Yu. Osipenko, Optimal Recovery of Analytic Functions, Nova Science Publ., Huntington, New York, 2000.

[11] M. S. Pinsker, Optimal filtration of square-integrable signals in Gaussian noise, Problems Inform. Transmission 16 (2) (1980) 120–133.

[12] L. Plaskota, Noisy Information and Computational Complexity, Cambridge University Press, Cambridge, 1996.

[13] S. Reshetov, Minimax risk over quadratically convex sets, J. Math. Sci., New York 167 (4) (2010) 537-542.

[14] J. F. Traub, H. Woźniakowski, A General Theory of Optimal Algorithms, Academic Press, New York, 1980.

[15] M. Wilczyński, Minimax estimation in linear regression with ellipsoidal constraints, J. Stat. Plann. Inference 137 (1) (2007) 79–86.

[16] M. Wilczyński, Minimax estimation over ellipsoids in l_2, Statistics 42 (2) (2008) 95–100.