DOI 10.1070/IM2014v078n06ABEH002724

On the best methods for recovering derivatives in Sobolev classes

G. G. Magaril-Il'yaev and K. Yu. Osipenko

Abstract. We construct the best (optimal) methods for recovering derivatives of functions in generalized Sobolev classes of functions on $\mathbb{R}^d$ provided that for every such function we know (exactly or approximately) its Fourier transform on an arbitrary measurable set $A \subset \mathbb{R}^d$. In both cases we construct families of optimal methods. These methods use only part of the information about the Fourier transform, and this part is subject to some filtration. We consider the problem of finding the best set for the recovery of a given derivative among all sets of a fixed measure.

Keywords: optimal recovery, Sobolev class, extremal problem, Fourier transform.

§ 1. Statement of the problems and results

Let d be a positive integer and F the Fourier transform in $L_2(\mathbb{R}^d)$. When $x(\,\cdot\,) \in L_2(\mathbb{R}^d)$, it is convenient to regard $Fx(\,\cdot\,)$ as a function on $\mathbb{R}^d$ with the Lebesgue measure divided by $(2\pi)^d$. The norm of a function $y(\,\cdot\,)$ in the space of square-integrable functions on $\mathbb{R}^d$ with such a measure is denoted by $\|y(\,\cdot\,)\|_{\widehat{L}_2(\mathbb{R}^d)}$, that is,

$$\|y(\,\cdot\,)\|_{\widehat{L}_2(\mathbb{R}^d)} = \biggl(\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} |y(\xi)|^2\,d\xi\biggr)^{1/2}.$$

For every $r>0$, the generalized Sobolev space (or the space of Bessel potentials) $\mathcal{H}_2^r(\mathbb{R}^d)$ is defined as the set of functions $x(\,\cdot\,) \in L_2(\mathbb{R}^d)$ such that

$$\|x(\,\cdot\,)\|_{\mathcal{H}_2^r(\mathbb{R}^d)} = \biggl(\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} (1+\|\xi\|^2)^r\,|(Fx)(\xi)|^2\,d\xi\biggr)^{1/2} < \infty,$$

where $\|\xi\|^2 = \xi_1^2 + \dots + \xi_d^2$. The corresponding generalized Sobolev class is the set

$$H_2^r(\mathbb{R}^d) = \bigl\{x(\,\cdot\,) \in \mathcal{H}_2^r(\mathbb{R}^d) \bigm| \|x(\,\cdot\,)\|_{\mathcal{H}_2^r(\mathbb{R}^d)} \leqslant 1\bigr\}.$$

For a positive integer r, a function $x(\,\cdot\,)$ (of the variables $t_1,\dots,t_d$) belongs to $\mathcal{H}_2^r(\mathbb{R}^d)$ if and only if all its generalized derivatives $\partial^{\alpha_1+\dots+\alpha_d}x/\partial t_1^{\alpha_1}\cdots\partial t_d^{\alpha_d}$

This paper was written with the financial support of RFBR (grants nos. 13-01-12447, 14-01-00456).

AMS 2010 *Mathematics Subject Classification.* 26D15, 42B10, 49K35, 90C47.

with $(\alpha_1,\dots,\alpha_d)\in\mathbb{Z}_+^d$ and $\alpha_1+\dots+\alpha_d\leqslant r$ belong to $L_2(\mathbb{R}^d)$. In this case, the norm

$$\sum_{\alpha_1+\dots+\alpha_d\leqslant r}\biggl\|\frac{\partial^{\alpha_1+\dots+\alpha_d}x(\,\cdot\,)}{\partial t_1^{\alpha_1}\cdots\partial t_d^{\alpha_d}}\biggr\|_{L_2(\mathbb{R}^d)}$$

is equivalent to the norm in $\mathcal{H}_2^r(\mathbb{R}^d)$. Thus, for positive integers r, $\mathcal{H}_2^r(\mathbb{R}^d)$ is the classical Sobolev space of functions on $\mathbb{R}^d$.

We now define fractional derivatives. For $\alpha=(\alpha_1,\dots,\alpha_d)\in\mathbb{R}_+^d$ and $\xi=(\xi_1,\dots,\xi_d)\in\mathbb{R}^d$ we put $(i\xi)^\alpha=(i\xi_1)^{\alpha_1}\cdots(i\xi_d)^{\alpha_d}$, where $(i\xi_j)^{\alpha_j}=|\xi_j|^{\alpha_j}\times\exp\bigl\{\frac12\pi i\operatorname{sign}\xi_j\bigr\}$, $j=1,\dots,d$ ($\operatorname{sign}0=0$, $0^0=1$), and let $\mathcal{E}^\alpha$ be the operator of multiplication by the function $\xi\mapsto(i\xi)^\alpha$ in $\widehat{L}_2(\mathbb{R}^d)$. If a function $x(\,\cdot\,)\in L_2(\mathbb{R}^d)$ is such that $(\mathcal{E}^\alpha\circ F)x(\,\cdot\,)\in\widehat{L}_2(\mathbb{R}^d)$, then the following function is well defined:

$$D^\alpha x(\,\cdot\,)=(F^{-1}\circ\mathcal{E}^\alpha\circ F)x(\,\cdot\,)\in L_2(\mathbb{R}^d),$$

where F^{-1} is the inverse Fourier transform. This function is called the α *th derivative* (in the sense of Weyl) of $x(\,\cdot\,)$. Clearly, if $x(\,\cdot\,)$ is sufficiently smooth and rapidly decreasing on $\mathbb{R}^d$ and $\alpha=(\alpha_1,\dots,\alpha_d)\in\mathbb{Z}_+^d$, then

$$D^\alpha x(t)=\frac{\partial x^{\alpha_1+\dots+\alpha_d}(t)}{\partial t_1^{\alpha_1}\cdots\partial t_d^{\alpha_d}}.$$

We are interested in questions which may informally be stated as follows.

1. Suppose that for every function $x(\,\cdot\,)\in H_2^r(\mathbb{R}^d)$ we know (exactly or approximately) its Fourier transform on some subset of $\mathbb{R}^d$. Then what is the best way to recover $D^\alpha x(\,\cdot\,)$ from this information?

2. Suppose that we can measure (exactly or approximately) the Fourier transforms of the functions $x(\,\cdot\,)\in H_2^r(\mathbb{R}^d)$ on any set of measure not exceeding some number $\sigma>0$. In other words, we can measure a fixed 'number of harmonics'. Which harmonics should be taken for the best recovery of $D^\alpha x(\,\cdot\,)$?

We now state the problems 1, 2 exactly. Suppose that A is an arbitrary measurable subset of $\mathbb{R}^d$ and, for every function $x(\,\cdot\,)\in H_2^r(\mathbb{R}^d)$, we know its Fourier transform on A either exactly or within accuracy $\delta>0$ in the metric of $\widehat{L}_2(A)$, that is, we know a function $y(\,\cdot\,)\in\widehat{L}_2(A)$ such that $\|(Fx)(\,\cdot\,)-y(\,\cdot\,)\|_{\widehat{L}_2(A)}\leqslant\delta$. From this information we want to recover $D^\alpha x(\,\cdot\,)$, $\alpha\in\mathbb{R}_+^d$, in the metric of $L_2(\mathbb{R}^d)$. This is understood in the following sense.

Let $I^\delta(A)\colon H_2^r(\mathbb{R}^d)\to\widehat{L}_2(A)$ be the map assigning to each function $x(\,\cdot\,)\in H_2^r(\mathbb{R}^d)$ the set $I^\delta(A)x(\,\cdot\,)=\{\,y(\,\cdot\,)\in\widehat{L}_2(A)\mid\|Fx(\,\cdot\,)-y(\,\cdot\,)\|_{\widehat{L}_2(A)}\leqslant\delta\}$ ($I^0(A)$ is the familiar map sending each function $x(\,\cdot\,)$ to the restriction $Fx(\,\cdot\,)|_A$ of the function $Fx(\,\cdot\,)$ to A). We denote the image of this map by $\operatorname{Im}I^\delta(A)$.

A method of recovery must associate with every function (observation) $y(\,\cdot\,)\in\operatorname{Im}I^\delta(A)$ a function in $L_2(\mathbb{R}^d)$ which is approximately equal to the αth derivative of the function in $H_2^r(\mathbb{R}^d)$. Thus every method is a map $\varphi\colon\operatorname{Im}I^\delta(A)\to L_2(\mathbb{R}^d)$. The error of a method is defined as

$$e\bigl(D^\alpha,H_2^r(\mathbb{R}^d),A,\delta,\varphi\bigr)=\sup_{\substack{x(\,\cdot\,)\in H_2^r(\mathbb{R}^d)\\ y(\,\cdot\,)\in I^\delta(A)}}\bigl\|D^\alpha x(\,\cdot\,)-\varphi(y(\,\cdot\,))(\,\cdot\,)\bigr\|_{L_2(\mathbb{R}^d)}.$$

When $\delta = 0$ this can be written in a shorter form:

$$e\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A, 0, \varphi\bigr) = \sup_{x(\,\cdot\,)\in H_2^r(\mathbb{R}^d)} \bigl\|D^\alpha x(\,\cdot\,) - \varphi(Fx(\,\cdot\,)|_A)(\,\cdot\,)\bigr\|_{L_2(\mathbb{R}^d)}.$$

We are interested in the quantity

$$E\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A, \delta\bigr) = \inf_{\varphi} e\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A, \delta, \varphi\bigr),$$

where the infimum is taken over all methods $\varphi\colon \operatorname{Im} I^\delta(A) \to L_2(\mathbb{R}^d)$. This quantity is called the *optimal recovery error*. We are also interested in the methods $\widehat{\varphi}$ at which the infimum is attained, that is,

$$E\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A, \delta\bigr) = e\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A, \delta, \widehat{\varphi}\bigr).$$

Such methods $\widehat{\varphi}$ are called *optimal recovery methods*.

An exact statement of problem 1 is to find the optimal recovery error and optimal recovery methods.

For every $\sigma > 0$ let $\mathcal{A}_\sigma$ be the family of all measurable subsets of $\mathbb{R}^d$ whose Lebesgue measure does not exceed σ. We are interested in the quantity

$$E_\sigma\bigl(D^\alpha, H_2^r(\mathbb{R}^d), \delta\bigr) = \inf_{A_\sigma \in \mathcal{A}_\sigma} E\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A_\sigma, \delta\bigr) \tag{1}$$

and in the sets at which the infimum is attained. Such sets are called *optimal sets*.

An exact statement of problem 2 is to find the quantity (1) and the optimal sets.

The original ideas underlying these problems date back to Kolmogorov, who introduced in [1] the notion of width, the quantity characterizing the best approximation of a class of functions by subspaces of a fixed dimension. The study of best quadratures on classes of functions began in the 1950s (the first investigations were those of Sard [2] and Nikol'skii [3]). In 1965 Smolyak [4] posed the general problem of the optimal recovery of a linear function on a class of elements from imprecise information about these elements. He proved that if this class is a convex centrally symmetric set, then there is a linear optimal method. Subsequently, the more general problem of the recovery of linear operators was posed, and the theory of optimal recovery underwent rapid development. One can get an impression of this from the surveys and monographs [5]–[11]. The optimal recovery problems studied in [12]–[17] are close to those considered in the present paper. We mention separately our paper [18], whose subject is the same problem as here but with A being the whole space $\mathbb{R}^d$.

Before stating our main results we introduce some definitions and notation.

Given any $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{Z}_+^d$ and $\xi = (\xi_1, \dots, \xi_d) \in \mathbb{R}^d$, we put

$$\overline{\alpha} = \sum_{j=1}^d \alpha_j, \qquad \alpha^\alpha = \prod_{j=1}^d \alpha_j^{\alpha_j}, \qquad |\xi|^\alpha = \prod_{j=1}^d |\xi_j|^{\alpha_j}.$$

Suppose that $0 < \overline{\alpha} < r$. We define a function $f(\,\cdot\,)$ on $\mathbb{R}^d$ by the formula

$$f(\xi) = \frac{|\xi|^{2\alpha}}{(1 + \|\xi\|^2)^r}, \qquad \xi \in \mathbb{R}^d.$$

Clearly, f is bounded and tends to zero as $\|\xi\| \to \infty$. A simple calculation shows that its maximal value

$$\widehat{\lambda} = \frac{\alpha^{\alpha}(r-\overline{\alpha})^{r-\overline{\alpha}}}{r^r}$$

is attained only at the points

$$\left(\pm\sqrt{\frac{\alpha_1}{r-\overline{\alpha}}},\dots,\pm\sqrt{\frac{\alpha_d}{r-\overline{\alpha}}}\,\right).$$

We put

$$\widehat{\delta} = \left(\frac{r-\overline{\alpha}}{r}\right)^{r/2}$$

and define a function h on the half-line $[0,\infty]$ by the formula

$$h(t) = \begin{cases} \dfrac{\alpha^{\alpha}}{r\,\overline{\alpha}^{\,\overline{\alpha}-1}}(1-t^{2/r})^{\overline{\alpha}-1}t^{2(1-\overline{\alpha}/r)}, & 0 \leqslant t \leqslant \widehat{\delta}, \\ \widehat{\lambda}, & t \geqslant \widehat{\delta}. \end{cases}$$

It is easy to see that $h(t)$ is strictly increasing on the closed interval $[0,\widehat{\delta}]$, $h(0)=0$ and $h(\widehat{\delta}) = \widehat{\lambda}$.

For every $\lambda \geqslant 0$ we define a set

$$\Omega_\lambda = \bigl\{\xi \in \mathbb{R}^d \mid f(\xi) \geqslant \lambda\bigr\}$$

and associate with every measurable subset A of $\mathbb{R}^d$ a number

$$\lambda(A) = \inf\bigl\{\lambda > 0 \mid \operatorname{mes}(A \cap \Omega_\lambda) = \operatorname{mes}\Omega_\lambda\bigr\}.$$

Clearly, $\operatorname{mes}(A\cap\Omega_{\widehat{\lambda}}) = \operatorname{mes}\Omega_{\widehat{\lambda}}$ since $\operatorname{mes}\Omega_{\widehat{\lambda}} = 0$. On the other hand, if A coincides a. e. with $\mathbb{R}^d$, then $\lambda(A) = 0$. Thus, $0 \leqslant \lambda(A) \leqslant \widehat{\lambda}$.

If $\delta = 0$ and $\lambda(A) = 0$, then we have complete information about the desired function, and the recovery problem becomes obvious. Therefore we do not consider the case $\delta + \lambda(A) = 0$ in what follows.

For every $\delta \geqslant 0$ and each measurable subset A of $\mathbb{R}^d$ with $\delta + \lambda(A) \neq 0$, we put

$$\Delta = \Delta(\delta, A) = \begin{cases} \delta, & 0 < \delta < \widehat{\delta}, \quad \lambda(A) \leqslant h(\delta), \\ h^{-1}(\lambda(A)), & 0 \leqslant \delta < \widehat{\delta}, \quad \lambda(A) > h(\delta), \\ \widehat{\delta}, & \delta \geqslant \widehat{\delta}, \end{cases}$$

and define the numbers

$$\lambda_1 = \lambda_1(\delta, A) = \frac{r}{\overline{\alpha}\Delta^2}(\widehat{\delta}^{\,2/r} - \Delta^{2/r})h(\Delta), \qquad \lambda_2 = \lambda_2(\delta, A) = h(\Delta).$$

Note that $h(\Delta) = \max\{\lambda(A), h(\delta)\}$.

Finally, we write $\langle\cdot,\cdot\rangle$ for the scalar product in $\mathbb{R}^d$.

Theorem 1. *Suppose that $\alpha \in \mathbb{R}^d_+$, $0 < \overline{\alpha} < r$, $\delta \geqslant 0$, A is a measurable subset of $\mathbb{R}^d$ and $\delta + \lambda(A) \neq 0$. Then $D^\alpha x(\,\cdot\,) \in L_2(\mathbb{R}^d)$ for every function $x(\,\cdot\,) \in \mathcal{H}^r_2(\mathbb{R}^d)$ and we have*

$$E(D^\alpha, H^r_2(\mathbb{R}^d), A, \delta) = \sqrt{\left(\frac{r\delta^2}{\overline{\alpha}\Delta^2}(\widehat{\delta}^{2/r} - \Delta^{2/r}) + 1\right)h(\Delta)}. \tag{2}$$

If $\delta = 0$, if $a(\,\cdot\,)$ is any measurable function on A satisfying

$$|a(\xi) - 1| \leqslant \sqrt{\lambda(A)}\,\frac{(1 + \|\xi\|^2)^{r/2}}{|\xi|^\alpha} \tag{3}$$

for a. e. $\xi \in A$, and if we define a method $\widehat{\varphi}_a$ by the following rule for a. e. $t \in \mathbb{R}^d$:

$$\widehat{\varphi}_a(Fx(\,\cdot\,)|_A)(t) = \frac{1}{(2\pi)^d}\int_A (i\xi)^\alpha a(\xi) Fx(\xi) e^{\langle \xi, t\rangle}\, d\xi,$$

then the method $\widehat{\varphi}_a$ is optimal.

If $\delta > 0$, if $a(\,\cdot\,)$ is any measurable function on A satisfying

$$\left|a(\xi) - \frac{\lambda_1}{\lambda_1 + \lambda_2(1 + \|\xi\|^2)^r}\right| \leqslant \frac{\sqrt{\lambda_1\lambda_2}(1 + \|\xi\|^2)^{r/2}}{|\xi|^\alpha(\lambda_1 + \lambda_2(1 + \|\xi\|^2)^r)}\sqrt{-|\xi|^{2\alpha} + \lambda_1 + \lambda_2(1 + \|\xi\|^2)^r} \tag{4}$$

for a. e. $\xi \in A$, and if we define a method $\widehat{\varphi}_a$ by the following rule for a. e. $t \in \mathbb{R}^d$:

$$\widehat{\varphi}_a(y(\,\cdot\,))(t) = \frac{1}{(2\pi)^d}\int_A (i\xi)^\alpha a(\xi) y(\xi) e^{\langle \xi, t\rangle}\, d\xi,$$

then the method $\widehat{\varphi}_a$ is optimal.

We now comment on Theorem 1.

1. If $\delta = 0$, then the expression (2) for the optimal recovery error implies that

$$E\bigl(D^\alpha, H^r_2(\mathbb{R}^d), A, 0\bigr) = \sqrt{\lambda(A)}.$$

Hence it suffices to know the Fourier transforms of functions in $H^r_2(\mathbb{R}^d)$ only on a set $A' \subset A$ (the inclusion is understood up to a set of measure zero) with $\lambda(A') = \lambda(A)$. The set $\Omega_{\lambda(A)}$ is minimal among such sets.

The optimal method is the αth derivative of the function whose Fourier transform vanishes outside A and is equal on A to the 'smoothing' of $Fx(\,\cdot\,)$ by means of $a(\,\cdot\,)$. The function $\xi \mapsto (i\xi)^\alpha a(\xi) Fx(\xi)$ belongs to $\widehat{L}_2(A)$ by (3). If it also belongs to $L_1(A)$ (for example, when the measure of A is finite), then the expression for the optimal method is the Fourier inversion formula. Otherwise the integral in the expression for the optimal method should be understood as the principal value for every $t \in \mathbb{R}^d$.

The function $f(\,\cdot\,)$ does not exceed $\lambda(A)$ on the set $A \setminus \Omega_{\lambda(A)}$. By (3), one can put $a(\,\cdot\,) = 0$ on this set and, therefore, it suffices to integrate only over $\Omega_{\lambda(A)}$.

If we put $a(\,\cdot\,) = 1$ on A, then (3) holds trivially. Hence one need not smooth the observed Fourier transform.

Finally, if $\lambda(A) = \widehat{\lambda}$, then $f(\xi) \leqslant \widehat{\lambda}$ for all $\xi \in \mathbb{R}^d$. The inequality (3) holds when $a(\,\cdot\,) = 0$ on A and, therefore, the zero method is optimal in this case.

To summarize, we see that the optimal method

$$\widehat{\varphi}(Fx(\,\cdot\,)|_A)(t) = \frac{1}{(2\pi)^d} \int_{\Omega_{\lambda(A)}} (i\xi)^\alpha Fx(\xi) e^{\langle \xi, t\rangle}\, d\xi$$

is the most 'reasonable' because it uses a minimal amount of information about the Fourier transform and requires no processing of this information. (Moreover, in the case when $\lambda(A) = \widehat{\lambda}$, the integral is taken over a set of measure zero and, therefore, $\widehat{\varphi} = 0$.)

2. If $\delta > 0$, then a straightforward (but quite routine) calculation shows that the optimal recovery error is a decreasing function of $\lambda(A)$ provided that $\lambda(A)$ decreases from $\widehat{\lambda}$ to $h(\delta)$. It follows easily from the definition of Δ that this error then stabilizes at the level

$$\sqrt{\biggl(\frac{r}{\overline{\alpha}}(\widehat{\delta}^{2/r} - \delta^{2/r}) + 1\biggr) h(\delta)}.$$

Thus the information on the Fourier transform outside a set with $\lambda(A) \leqslant h(\delta)$ turns out to be redundant. The set $\Omega_{h(\delta)}$ is minimal among such sets.

In Theorem 1 we represent a family of optimal methods, each of which is the αth derivative of a function whose Fourier transform vanishes outside A and coincides on A with the 'smoothing' of $y(\,\cdot\,)$ by means of $a(\,\cdot\,)$.

One can put $a(\,\cdot\,) = 0$ on $A \setminus \Omega_{h(\Delta)}$ (we recall that $h(\Delta) = \max\{\lambda(A), h(\delta)\}$). Indeed, in this case the inequality (23) in § 2, which is equivalent to (4), implies that we must have $f(\xi) \leqslant \lambda_2$ on $A \setminus \Omega_{h(\Delta)}$. This inequality does indeed hold because $\lambda_2 = h(\Delta)$. Thus, for every $a(\,\cdot\,)$ satisfying (4), the most economical optimal method is given by

$$\widehat{\varphi}_a(y(\,\cdot\,))(t) = \frac{1}{(2\pi)^d} \int_{\Omega_{h(\Delta)}} (i\xi)^\alpha a(\xi) y(\xi) e^{\langle \xi, t\rangle}\, d\xi.$$

We also present explicitly the optimal method corresponding to a function $a(\,\cdot\,)$ for which the left-hand side of (4) is equal to zero:

$$\widehat{\varphi}(y(\,\cdot\,))(t) \\ = \frac{1}{(2\pi)^d} \int_{\Omega_{h(\Delta)}} (i\xi)^\alpha \biggl(1 + \frac{\overline{\alpha}\Delta^2}{r} \frac{1}{\widehat{\delta}^{2/r} - \Delta^{2/r}} (1 + \|\xi\|^2)^r\biggr)^{-1} y(\xi) e^{\langle \xi, t\rangle} d\xi.$$

Before stating our next theorem, we give some definitions. Clearly, the function $m\colon \lambda \mapsto \operatorname{mes}\Omega_\lambda$ is monotone decreasing on $(0, \widehat{\lambda}]$, $m(\lambda) \to \infty$ as $\lambda \to 0$ and $m(\widehat{\lambda}) = 0$. For every $\sigma > 0$ let $\lambda(\sigma)$ be the unique solution of the equation $m(\lambda) = \sigma$.

Theorem 2. *Suppose that $\alpha \in \mathbb{R}^d_+$, $0 < \overline{\alpha} < r$, $\delta \geqslant 0$, $\sigma > 0$ and $\lambda(\sigma, \delta) = \max(\lambda(\sigma), h(\delta))$. Then every set that coincides with $\Omega_{\lambda(\sigma,\delta)}$ up to a set of measure zero is optimal.*

We note that if $\delta = 0$, then $\lambda(\sigma,\delta) = \lambda(\sigma)$ because $h(0) = 0$. Therefore $\Omega_{\lambda(\sigma)}$ is an optimal set, and the optimal recovery error $\sqrt{\lambda(\sigma)}$ becomes smaller as σ increases.

If $\delta > 0$ and $\sigma > 0$ are such that $\lambda(\sigma) \leqslant h(\delta)$, then the information on the Fourier transform outside $\Omega_{h(\delta)}$ turns out to be redundant since the optimal recovery error does not decrease.

§ 2. Proofs

Proof of Theorem 1. We first claim that if $x(\,\cdot\,) \in \mathcal{H}_2^r(\mathbb{R}^d)$, then $D^\alpha x(\,\cdot\,) \in L_2(\mathbb{R}^d)$. Indeed, by the definitions of $\widehat{\lambda}$ and $\mathcal{H}_2^r(\mathbb{R}^d)$ we have

$$\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} |\xi|^{2\alpha}\,|Fx(\xi)|^2\,d\xi = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \frac{|\xi|^{2\alpha}}{(1+\|\xi\|^2)^r}(1+\|\xi\|^2)^r\,|Fx(\xi)|^2\,d\xi$$
$$\leqslant \widehat{\lambda}\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}(1+\|\xi\|^2)^r\,|Fx(\xi)|^2\,d\xi < \infty.$$

Using the definition of the αth derivative and Plancherel's theorem, we get

$$\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} |\xi|^{2\alpha}\,|Fx(\xi)|^2\,d\xi = \int_{\mathbb{R}^d} |D^\alpha x(t)|^2\,dt,$$

that is, $D^\alpha x(\,\cdot\,) \in L_2(\mathbb{R}^d)$.

We now obtain a lower bound for the optimal recovery error $E(D^\alpha, H_2^r(\mathbb{R}^d), A, \delta)$. We claim that this error is not less than the value of the extremal problem

$$\|D^\alpha x(\,\cdot\,)\|_{L_2(\mathbb{R}^d)} \to \max, \quad \|Fx(\,\cdot\,)\|_{\widehat{L}_2(A)} \leqslant \delta, \quad \|x(\,\cdot\,)\|_{\mathcal{H}_2^r(\mathbb{R}^d)} \leqslant 1, \tag{5}$$

that is, the supremum of the functional maximized under these constraints. (If $\delta = 0$, then the first constraint takes the form $Fx(\,\cdot\,) = 0$ for a. e. $\xi \in A$.)

Indeed, let $x_0(\,\cdot\,)$ be an admissible function for (5) (that is, $x_0(\,\cdot\,)$ satisfies the constraints). Then, clearly, the function $-x_0(\,\cdot\,)$ is also admissible and, for every $\varphi\colon \widehat{L}_2(A) \to L_2(\mathbb{R}^d)$ (where $\varphi(0)(\,\cdot\,)$ is the value of φ on the zero function) we have

$$2\|D^\alpha x_0(\,\cdot\,)\|_{L_2(\mathbb{R}^d)} \leqslant \big\|D^\alpha x_0(\,\cdot\,) - \varphi(0)(\,\cdot\,)\big\|_{L_2(\mathbb{R}^d)} + \big\|D^\alpha(-x_0)(\,\cdot\,) - \varphi(0)(\,\cdot\,)\big\|_{L_2(\mathbb{R}^d)}$$
$$\leqslant 2 \sup_{\substack{x(\cdot)\in H_2^r(\mathbb{R}^d)\\ \|Fx(\cdot)\|_{\widehat{L}_2(A)}\leqslant\delta}} \big\|D^\alpha x(\,\cdot\,) - \varphi(0)(\,\cdot\,)\big\|_{L_2(\mathbb{R}^d)}$$
$$\leqslant 2 \sup_{\substack{x(\cdot)\in H_2^r(\mathbb{R}^d),\,y(\cdot)\in\widehat{L}_2(A)\\ \|Fx(\cdot)-y(\cdot)\|_{\widehat{L}_2(A)}\leqslant\delta}} \big\|D^\alpha x(\,\cdot\,) - \varphi(y(\,\cdot\,))(\,\cdot\,)\big\|_{L_2(\mathbb{R}^d)}.$$

Taking the supremum of the left-hand side over all admissible functions for (5) and the infimum of the right-hand side over all methods φ, we get the desired result.

We now bound the value of the problem (5) from below. To do this, it is convenient to rewrite the problem in terms of Fourier images. By Plancherel's theorem, the squared value of (5) is equal to the value of the problem

$$\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}|\xi|^{2\alpha}\,|Fx(\xi)|^2\,d\xi\to\max,\qquad \frac{1}{(2\pi)^d}\int_A|Fx(\xi)|^2\,d\xi\leqslant\delta^2,$$
$$\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}(1+\|\xi\|^2)^r\,|Fx(\xi)|^2\,d\xi\leqslant 1. \tag{6}$$

When $\delta=0$, the first constraint in this problem is given by $Fx(\,\cdot\,)=0$ for a. e. $\xi\in A$.

We consider some cases separately.

Case 1: $\delta\geqslant\widehat{\delta}$. We put

$$\widehat{\xi}=\biggl(\sqrt{\frac{\alpha_1}{r-\overline{\alpha}}},\dots,\sqrt{\frac{\alpha_d}{r-\overline{\alpha}}}\,\biggr),$$

where $\alpha=(\alpha_1,\dots,\alpha_d)$. For every $\varepsilon>0$ we introduce the notation

$$\widehat{\xi}_\varepsilon=\biggl(1+\frac{\varepsilon}{\|\widehat{\xi}\|}\biggr)\widehat{\xi},$$

consider the ball $B_\varepsilon=\{\xi\in\mathbb{R}^d\mid\|\xi-\widehat{\xi}_\varepsilon\|\leqslant\varepsilon\}$ and define the following functions on $\mathbb{R}^d$:

$$z_\varepsilon(\xi)=\begin{cases}(2\pi)^{d/2}\biggl(\displaystyle\int_{B_\varepsilon}(1+\|\eta\|^2)^r\,d\eta\biggr)^{-1/2}, & \xi\in B_\varepsilon,\\ 0, & \xi\notin B_\varepsilon.\end{cases}$$

Clearly, $z_\varepsilon(\,\cdot\,)\in\widehat{L}_2(\mathbb{R}^d)$. We claim that the functions $x_\varepsilon(\,\cdot\,)=F^{-1}z_\varepsilon(\,\cdot\,)$ are admissible in the problem (6).

Indeed, the second requirement in (6) obviously holds. If $A\cap B_\varepsilon=\varnothing$, then the first requirement holds for trivial reasons.

Suppose that $A\cap B_\varepsilon\neq\varnothing$. We easily verify that $(1+\|\widehat{\xi}\|^2)^{-r}=\widehat{\delta}^{\,2}$ and $\|\xi\|\geqslant\|\widehat{\xi}\|$ for all $\xi\in B_\varepsilon$. Therefore we have

$$\frac{1}{(2\pi)^d}\int_A|Fx_\varepsilon(\xi)|^2\,d\xi=\frac{1}{(2\pi)^d}\int_{A\cap B_\varepsilon}|Fx_\varepsilon(\xi)|^2\,d\xi\leqslant\biggl(\int_{B_\varepsilon}(1+\|\eta\|^2)^r\,d\eta\biggr)^{-1}\operatorname{mes}B_\varepsilon$$
$$\leqslant\biggl(\int_{B_\varepsilon}(1+\|\widehat{\xi}\|^2)^r\,d\eta\biggr)^{-1}\operatorname{mes}B_\varepsilon=(1+\|\widehat{\xi}\,\|^2)^{-r}=\widehat{\delta}^{\,2}\leqslant\delta^2.$$

Thus, for every $\varepsilon>0$ the function $x_\varepsilon(\,\cdot\,)$ is admissible in the problem (5) and, therefore, the value of this problem is not less than

$$\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}|\xi|^{2\alpha}\,|Fx_\varepsilon(\xi)|^2\,d\xi=\biggl(\int_{B_\varepsilon}(1+\|\eta\|^2)^r\,d\eta\biggr)^{-1}\int_{B_\varepsilon}|\xi|^{2\alpha}\,d\xi.$$

As $\varepsilon\to 0$, this quantity tends (by the mean value theorem) to the quantity

$$f(\widehat{\xi}\,)=\frac{|\widehat{\xi}\,|^{2\alpha}}{(1+\|\widehat{\xi}\|^2)^r},$$

which is equal to $\widehat{\lambda}$ since the maximum of f is attained at the point $\widehat{\xi}$ (this was mentioned before statement of the theorem).

Since the squared optimal recovery error is not less than the value of (5), we obtain the following estimate for $\delta \geqslant \widehat{\delta}$:

$$E\big(D^\alpha, H_2^r(\mathbb{R}^d), A, \delta\big) \geqslant \sqrt{\widehat{\lambda}}. \tag{7}$$

The right-hand side of (7) coincides with the optimal recovery error indicated in the theorem because in the case considered we have $\Delta = \widehat{\delta}$ and $h(\widehat{\delta}) = \widehat{\lambda}$.

Case 2: $0 < \delta < \widehat{\delta}$, $\lambda(A) \leqslant h(\delta)$. Put

$$\widetilde{\xi} = \left(\sqrt{\frac{\alpha_1(1-\delta^{2/r})}{\overline{\alpha}\delta^{2/r}}}, \ldots, \sqrt{\frac{\alpha_d(1-\delta^{2/r})}{\overline{\alpha}\delta^{2/r}}}\,\right).$$

For every $\varepsilon > 0$ we introduce the notation

$$\widetilde{\xi}_\varepsilon = \left(1 - \frac{\varepsilon}{\|\widetilde{\xi}\|}\right)\widetilde{\xi}$$

and consider the ball $\widetilde{B}_\varepsilon = \{\xi \in \mathbb{R}^d \mid \|\xi - \widetilde{\xi}_\varepsilon\| \leqslant \varepsilon\}$.

Since $\delta < \widehat{\delta}$ and, therefore, $r(1-\delta^{2/r})/\overline{\alpha} > 1$, we have

$$\begin{aligned} f(\widetilde{\xi}) = \frac{|\widetilde{\xi}|^{2\alpha}}{(1+\|\widetilde{\xi}\|^2)^r} &= \frac{\alpha^\alpha}{r\overline{\alpha}^{\overline{\alpha}-1}}(1-\delta^{2/r})^{\overline{\alpha}-1}\delta^{2(1-\overline{\alpha}/r)} \\ &= \frac{r}{\overline{\alpha}}(1-\delta^{2/r})h(\delta) > h(\delta) \geqslant \lambda(A). \end{aligned}$$

It follows that $\widetilde{\xi} \in \operatorname{int}\Omega_{\lambda(A)}$, whence $B_\varepsilon \subset \operatorname{int}\Omega_{\lambda(A)}$ for sufficiently small ε and, therefore, $\operatorname{mes}(A \cap \widetilde{B}_\varepsilon) = \operatorname{mes}\widetilde{B}_\varepsilon$ for such ε. We define the following functions on $\mathbb{R}^d$:

$$z_\varepsilon(\xi) = \begin{cases} (2\pi)^{d/2}\dfrac{\delta}{\sqrt{\operatorname{mes}\widetilde{B}_\varepsilon}}, & \xi \in \widetilde{B}_\varepsilon, \\ 0, & \xi \notin \widetilde{B}_\varepsilon. \end{cases}$$

Clearly, $z_\varepsilon(\,\cdot\,) \in \widehat{L}_2(\mathbb{R}^d)$. We claim that the functions $x_\varepsilon(\,\cdot\,) = F^{-1}z_\varepsilon(\,\cdot\,)$ are admissible in the problem (6). The first requirement obviously holds. We easily verify that $\|\xi\| \leqslant \|\widetilde{\xi}\|$ if $\xi \in \widetilde{B}_\varepsilon$ and $\delta^2(1+\|\widetilde{\xi}\|^2)^r = 1$. Hence

$$\begin{aligned} \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}(1+\|\xi\|^2)^r\,|Fx_\varepsilon(\xi)|^2\,d\xi &= \frac{\delta^2}{\operatorname{mes}\widetilde{B}_\varepsilon}\int_{\widetilde{B}_\varepsilon}(1+\|\xi\|^2)^r\,d\xi \\ &\leqslant \frac{\delta^2}{\operatorname{mes}\widetilde{B}_\varepsilon}(1+\|\widetilde{\xi}\,\|^2)^r\operatorname{mes}\widetilde{B}_\varepsilon = \delta^2(1+\|\widetilde{\xi}\,\|^2)^r = 1. \end{aligned}$$

Thus the functions $x_\varepsilon(\,\cdot\,)$ are admissible in the problem (5) and, therefore, the value of this problem is not less than

$$\begin{aligned} \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}|\xi|^{2\alpha}\,|Fx_\varepsilon(\xi)|^2\,d\xi &= \frac{\delta^2}{\operatorname{mes}\widetilde{B}_\varepsilon}\int_{\widetilde{B}_\varepsilon}|\xi|^{2\alpha}\,d\xi \\ &= \frac{1}{(1+\|\widetilde{\xi}\|^2)^r\operatorname{mes}\widetilde{B}_\varepsilon}\int_{\widetilde{B}_\varepsilon}|\xi|^{2\alpha}\,d\xi \end{aligned}$$

for all sufficiently small $\varepsilon > 0$. As $\varepsilon \to 0$, this quantity tends (by the mean value theorem) to the number

$$\frac{|\widetilde{\xi}|^{2\alpha}}{(1+\|\widetilde{\xi}\|^2)^r} = \frac{\alpha^\alpha}{\overline{\alpha}^{\overline{\alpha}}} \frac{(1-\delta^{2/r})^{\overline{\alpha}}}{\delta^{2\overline{\alpha}/r}} \delta^2 = \left(\frac{r}{\overline{\alpha}}(\widehat{\delta}^{\,2/r} - \delta^{2/r}) + 1\right) h(\delta).$$

Arguing as above, we get the following estimate in the case considered:

$$E\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A, \delta\bigr) \geqslant \sqrt{\left(\frac{r}{\overline{\alpha}}(\widehat{\delta}^{2/r} - \delta^{2/r}) + 1\right) h(\delta)}. \tag{8}$$

The right-hand side coincides with the optimal recovery error indicated in the theorem because $\Delta = \delta$ in this case.

Case 3: $0 \leqslant \delta < \widehat{\delta}$, $\lambda(A) > h(\delta)$. We first assume that $\lambda(A) = \widehat{\lambda}$. Then $\operatorname{mes}(A \cap \Omega_\lambda) < \operatorname{mes} \Omega_\lambda$ for all λ, $0 < \lambda < \widehat{\lambda}$. Hence, for such λ, we have

$$\operatorname{mes}(\Omega_\lambda \cap (\mathbb{R}^d \setminus A)) = \operatorname{mes}(\Omega_\lambda \setminus A) = \operatorname{mes} \Omega_\lambda - \operatorname{mes}(A \cap \Omega_\lambda) > 0.$$

We put $G_\lambda = \Omega_\lambda \cap (\mathbb{R}^d \setminus A)$ and define the following functions on $\mathbb{R}^d$:

$$z_\lambda(\xi) = \begin{cases} (2\pi)^{d/2}\left(\displaystyle\int_{G_\lambda} (1+\|\eta\|^2)^r \, d\eta\right)^{-1/2}, & \xi \in G_\lambda, \\ 0, & \xi \notin G_\lambda. \end{cases}$$

Clearly, $z_\lambda(\,\cdot\,) \in \widehat{L}_2(\mathbb{R}^d)$. We put $x_\lambda(\,\cdot\,) = F^{-1} z_\lambda(\,\cdot\,)$. It is easily verified that the functions $x_\lambda(\,\cdot\,)$ are admissible in the problem (6). For all indicated values of λ (taking into account that $f(\xi) \geqslant \lambda$ when $\xi \in \Omega_\lambda$), we have

$$\begin{aligned} \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} |\xi|^{2\alpha} \, |F x_\lambda(\xi)|^2 \, d\xi &= \left(\int_{G_\lambda} (1+\|\eta\|^2)^r \, d\eta\right)^{-1} \int_{G_\lambda} |\xi|^{2\alpha} \, d\xi \\ &= \left(\int_{G_\lambda} (1+\|\eta\|^2)^r \, d\eta\right)^{-1} \int_{G_\lambda} \frac{|\xi|^{2\alpha}}{(1+\|\xi\|^2)^r} (1+\|\xi\|^2)^r \, d\xi \\ &\geqslant \left(\int_{G_\lambda} (1+\|\eta\|^2)^r \, d\eta\right)^{-1} \lambda \int_{G_\lambda} (1+\|\xi\|^2)^r \, d\xi = \lambda, \end{aligned}$$

whence the value of (6) is not less than λ. Letting $\lambda \to \widehat{\lambda}$, we see that the value of this problem is not less than $\widehat{\lambda}$.

Again, since the squared optimal recovery error is not less than the value of (6), we get the following estimate when $\lambda(A) = \widehat{\lambda}$:

$$E\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A, \delta\bigr) \geqslant \sqrt{\widehat{\lambda}}, \tag{9}$$

where the quantity on the right again coincides with the optimal recovery error indicated in the theorem because $\Delta = \widehat{\delta}$ in this case.

We now assume that $\lambda(A) < \widehat{\lambda}$. First, we claim that the measure of the set

$$F_\varepsilon = (\mathbb{R}^d \setminus A) \cap (\Omega_{\lambda(A)-\varepsilon} \setminus \Omega_{\lambda(A)})$$

is positive for all ϵ, $0 < \varepsilon < \lambda(A)$. Indeed, suppose that $\operatorname{mes} F_\varepsilon = 0$ for some ε. Since $\Omega_{\lambda(A)} \subset \Omega_{\lambda(A)-\varepsilon}$, we get

$$\begin{aligned}\operatorname{mes}\bigl(A \cap (\Omega_{\lambda(A)-\varepsilon} \setminus \Omega_{\lambda(A)})\bigr) &= \operatorname{mes}(\Omega_{\lambda(A)-\varepsilon} \setminus \Omega_\lambda)\\ &= \operatorname{mes}\Omega_{\lambda(A)-\varepsilon} - \operatorname{mes}(\Omega_{\lambda(A)} \cap \Omega_{\lambda(A)-\varepsilon}) = \operatorname{mes}\Omega_{\lambda(A)-\varepsilon} - \operatorname{mes}\Omega_{\lambda(A)}.\end{aligned}$$

On the other hand,

$$\begin{aligned}\operatorname{mes}\bigl(A \cap (\Omega_{\lambda(A)-\varepsilon} \setminus \Omega_{\lambda(A)})\bigr) &= \operatorname{mes}\bigl((A \cap \Omega_{\lambda(A)-\varepsilon}) \setminus \Omega_{\lambda(A)}\bigr)\\ &= \operatorname{mes}(A \cap \Omega_{\lambda(A)-\varepsilon}) - \operatorname{mes}(A \cap \Omega_{\lambda(A)}).\end{aligned}$$

We easily see from the definition of $\lambda(A)$ that $\operatorname{mes}(A \cap \Omega_{\lambda(A)}) = \operatorname{mes}\Omega_{\lambda(A)}$. Then the expressions above yield that $\operatorname{mes}(A \cap \Omega_{\lambda(A)-\varepsilon}) = \operatorname{mes}\Omega_{\lambda(A)-\varepsilon}$ contrary to the definition of $\lambda(A)$. Thus $\operatorname{mes} F_\varepsilon \neq 0$ for $0 < \varepsilon < \lambda(A)$.

Assume first that $\delta = 0$. For ε as above we define the following functions on $\mathbb{R}^d$:

$$z_\varepsilon(\xi) = \begin{cases} (2\pi)^{d/2}\left(\displaystyle\int_{F_\varepsilon} (1 + \|\eta\|^2)^r \, d\eta\right)^{-1/2}, & \xi \in F_\varepsilon,\\ 0, & \xi \notin F_\varepsilon. \end{cases}$$

Clearly, $z_\varepsilon(\,\cdot\,) \in \widehat{L}_2(\mathbb{R}^d)$. We put $x_\varepsilon(\,\cdot\,) = F^{-1}z_\varepsilon(\,\cdot\,)$ and verify in an elementary manner that the functions $x_\varepsilon(\,\cdot\,)$ are admissible in the problem (6). The same argument as above shows that

$$\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} |\xi|^{2\alpha}\, |Fx_\varepsilon(\xi)|^2 \, d\xi \geqslant \lambda(A) - \varepsilon,$$

whence it follows (as above) that

$$E\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A, 0\bigr) \geqslant \sqrt{\lambda(A)}. \tag{10}$$

The right-hand side of (10) coincides with the optimal recovery error in the statement of the theorem because $\delta = 0$ and $h(\Delta) = \lambda(A)$ in the case considered.

We now assume that $\delta > 0$. Put

$$\xi' = \left(\sqrt{\frac{\alpha_1(1 - \Delta^{2/r})}{\overline{\alpha}\Delta^{2/r}}}, \ldots, \sqrt{\frac{\alpha_d(1 - \Delta^{2/r})}{\overline{\alpha}\,\Delta^{2/r}}}\ \right).$$

For every $\varepsilon > 0$ we introduce the notation

$$\xi'_\varepsilon = \left(1 - \frac{\varepsilon}{\|\xi'\|}\right)\xi'$$

and consider the ball $B'_\varepsilon = \{\xi \in \mathbb{R}^d \mid \|\xi - \xi'_\varepsilon\| \leqslant \varepsilon\}$.

Since $\lambda(A) < \widehat{\lambda}$, we have $\Delta = h^{-1}(\lambda(A)) < h^{-1}(\widehat{\lambda}) = \widehat{\delta}$. As in Case 2, it follows that

$$\begin{aligned}f(\xi') = \frac{|\xi'|^{2\alpha}}{(1 + \|\xi'\|^2)^r} &= \frac{\alpha^\alpha}{r\overline{\alpha}^{\,\overline{\alpha}-1}}(1 - \Delta^{2/r})^{\overline{\alpha}-1}\Delta^{2(1-\overline{\alpha}/r)}\\ &= \frac{r}{\overline{\alpha}}(1 - \Delta^{2/r})h(\Delta) > h(\Delta) = \lambda(A).\end{aligned}$$

This means that $\xi' \in \operatorname{int}\Omega_{\lambda(A)}$. Hence, for sufficiently small $\varepsilon > 0$, the ball B'_ε lies in $\Omega_{\lambda(A)}$ and, therefore, $B'_\varepsilon \cap F_\varepsilon = \varnothing$.

We easily verify that $\|\xi\| \leqslant \|\xi'\|$ when $\xi \in B'_\varepsilon$ and $(1+\|\xi'\|^2)^r = 1/\Delta^2$. Since $\lambda(A) > h(\delta)$, we further have $\Delta = h^{-1}(\lambda(A)) > \delta$ and, therefore,

$$\frac{\delta^2}{\operatorname{mes} B'_\varepsilon}\int_{B'_\varepsilon}(1+\|\xi\|^2)^r\,d\xi \leqslant \frac{\delta^2}{\operatorname{mes} B'_\varepsilon}\int_{B'_\varepsilon}(1+\|\xi'\|^2)^r\,d\xi = \delta^2(1+\|\xi'\|^2)^r = \frac{\delta^2}{\Delta^2} < 1.$$

We denote the leftmost expression by C_ε and, with ε as above, define the following functions on $\mathbb{R}^d$:

$$z_\varepsilon(\xi) = \begin{cases} (2\pi)^{d/2}\dfrac{\delta}{\sqrt{\operatorname{mes} B'_\varepsilon}}, & \xi \in B'_\varepsilon, \\ (2\pi)^{d/2}\sqrt{1-C_\varepsilon}\Big(\displaystyle\int_{F_\varepsilon}(1+\|\eta\|^2)^r\,d\eta\Big)^{-1/2}, & \xi \in F_\varepsilon, \\ 0, & \xi \notin B'_\varepsilon \cup F_\varepsilon. \end{cases}$$

Clearly, $z_\varepsilon(\,\cdot\,) \in \widehat{L}_2(\mathbb{R}^d)$. We claim that the functions $x_\varepsilon(\,\cdot\,) = F^{-1}z_\varepsilon(\,\cdot\,)$ are admissible in the problem (6).

Indeed, since $B'_\varepsilon \subset \Omega_{\lambda(A)}$, we have $\operatorname{mes} B'_\varepsilon = \operatorname{mes}(A \cap B'_\varepsilon)$ and, therefore,

$$\frac{1}{(2\pi)^d}\int_A |Fx_\varepsilon(\xi)|^2\,d\xi = \frac{\delta^2}{\operatorname{mes} B'_\varepsilon}\int_{A\cap B'_\varepsilon} d\xi = \frac{\delta^2}{\operatorname{mes} B'_\varepsilon}\operatorname{mes}(A\cap B'_\varepsilon) = \delta^2.$$

We further have

$$\begin{aligned}\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}(1+\|\xi\|^2)^r\,|Fx_\varepsilon(\xi)|^2\,d\xi &= \frac{\delta^2}{\operatorname{mes} B'_\varepsilon}\int_{B'_\varepsilon}(1+\|\xi\|^2)^r\,d\xi \\ &+ (1-C_\varepsilon)\Big(\int_{F_\varepsilon}(1+\|\eta\|^2)^r\,d\eta\Big)^{-1}\int_{F_\varepsilon}(1+\|\xi\|^2)^r\,d\xi = C_\varepsilon + 1 - C_\varepsilon = 1,\end{aligned}$$

whence the functions $x_\varepsilon(\,\cdot\,)$ with sufficiently small ε are admissible in the problem (6). Then, for every such ε, the value of this problem is not less than

$$\begin{aligned}\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}|\xi|^{2\alpha}\,|Fx_\varepsilon(\xi)|^2\,d\xi &= \frac{\delta^2}{\operatorname{mes} B'_\varepsilon}\int_{B'_\varepsilon}|\xi|^{2\alpha}\,d\xi \\ &+ (1-C_\varepsilon)\Big(\int_{F_\varepsilon}(1+\|\eta\|^2)^r\,d\eta\Big)^{-1}\int_{F_\varepsilon}|\xi|^{2\alpha}\,d\xi. \qquad (11)\end{aligned}$$

Proceeding as above, we see that the product of the last two factors in the second term of (11) is not less than $\lambda(A) - \varepsilon$. Estimating the first term and C_ε by the mean value theorem, we see that, as $\varepsilon \to 0$, the right-hand side of (11) tends to

$$\begin{aligned}\delta^2|\xi'|^{2\alpha} + \big(1-\delta^2(1+\|\xi'\|^2)^r\big)\lambda(A) &= \frac{\alpha^\alpha}{\overline{\alpha}^{\overline{\alpha}}}\,\frac{(1-\Delta^{2/r})^{\overline{\alpha}}}{\Delta^{2\overline{\alpha}/r}}\,\delta^2 \\ &+ \Big(1-\frac{\delta^2}{\Delta^2}\Big)\lambda(A) = \Big(\frac{r\delta^2}{\overline{\alpha}\Delta^2}(\widehat{\delta}^{\,2/r} - \Delta^{2/r}) + 1\Big)\lambda(A).\end{aligned}$$

Hence, when $0 \leqslant \delta < \widehat{\delta}$ and $\lambda(A) > h(\delta)$, we get the estimate

$$E\big(D^\alpha, H_2^r(\mathbb{R}^d), A, \delta\big) \geqslant \sqrt{\bigg(\frac{r\delta^2}{\overline{\alpha}\Delta^2}(\widehat{\delta}^{2/r} - \Delta^{2/r}) + 1\bigg)\lambda(A)}, \tag{12}$$

whose right-hand side coincides with the optimal recovery error in the statement of the theorem because $h(\Delta) = \lambda(A)$ in this case.

Thus, for all $\delta \geqslant 0$ and all measurable sets $A \subset \mathbb{R}^d$ with $\delta + \lambda(A) \neq 0$, we have obtained a lower bound for the optimal recovery error (see (7)–(10) and (12)), which coincides with the value of the optimal recovery error given in the statement of the theorem. We now obtain an upper bound for this quantity and construct the optimal methods.

We fix $\delta \geqslant 0$ and $A \subset \mathbb{R}^d$ with $\delta + \lambda(A) \neq 0$. The optimality of a method $\varphi\colon \operatorname{Im} I^\alpha(A) \to L_2(\mathbb{R}^d)$ means that its error, that is, the value of the problem

$$\begin{aligned} \big\|D^\alpha x(\,\cdot\,) - \varphi(y(\,\cdot\,))(\,\cdot\,)\big\|_{L_2(\mathbb{R}^d)} \to \max, \qquad x(\,\cdot\,) \in H_2^r(\mathbb{R}^d),\\ \big\|Fx(\,\cdot\,) - y(\,\cdot\,)\big\|_{\widehat{L}_2(A)} \leqslant \delta, \qquad y(\,\cdot\,) \in \widehat{L}_2(A), \end{aligned} \tag{13}$$

coincides with $E(D^\alpha, H_2^r(\mathbb{R}^d), A, \delta)$.

When $\delta = 0$, problem (13) can be rewritten as

$$\big\|D^\alpha x(\,\cdot\,) - \varphi(Fx(\,\cdot\,)|_A)(\,\cdot\,)\big\|_{L_2(\mathbb{R}^d)} \to \max, \qquad x(\,\cdot\,) \in H_2^r(\mathbb{R}^d). \tag{14}$$

We consider some cases separately.

Case (a): $\delta = 0$. Since the map $x(\,\cdot\,) \mapsto D^\alpha x(\,\cdot\,)$ in the Fourier images is the multiplication of the function $\xi \mapsto (i\xi)^\alpha$ by $Fx(\,\cdot\,)$, it is natural to search for optimal methods among such maps. For every measurable function $a(\,\cdot\,)$ on A with $a(\,\cdot\,)\sqrt{f(\,\cdot\,)} \in L_\infty(A)$ we consider the map $\varphi_a\colon \operatorname{Im} I^0(A) \to L_2(\mathbb{R}^d)$, which acts on Fourier images by the rule $F\varphi_a(y(\,\cdot\,))(\xi) = (i\xi)^\alpha \widetilde{a}(\xi)\widetilde{y}(\xi)$ for a. e. $\xi \in \mathbb{R}^d$, where $\widetilde{a}(\,\cdot\,) = a(\,\cdot\,)$, $\widetilde{y}(\,\cdot\,) = y(\,\cdot\,)$ on A and $\widetilde{a}(\,\cdot\,) = 0$, $\widetilde{y}(\,\cdot\,) = 0$ outside A. This map is well defined because if $y(\,\cdot\,) \in \operatorname{Im} I^0(A)$, then $y(\,\cdot\,) = Fx(\,\cdot\,)|_A$ for some $x(\,\cdot\,) \in \mathcal{H}_2^r(\mathbb{R}^d)$, and then

$$\begin{aligned} \frac{1}{(2\pi)^d}\int_A \big|F\varphi_a(y(\,\cdot\,))(\xi)\big|^2\,d\xi &= \frac{1}{(2\pi)^d}\int_A |\xi|^{2\alpha}\,|a(\xi)|^2\,|Fx(\xi)|^2\,d\xi \\ &= \frac{1}{(2\pi)^d}\int_A |a(\xi)|^2 \frac{|\xi|^{2\alpha}}{(1+\|\xi\|^2)^r}(1+\|\xi\|^2)^r\,|Fx(\xi)|^2\,d\xi \\ &\leqslant \big\|\,|a(\,\cdot\,)|^2 f(\,\cdot\,)\big\|_{L_\infty(A)} \frac{1}{(2\pi)^d}\int_A (1+\|\xi\|^2)^r\,|Fx(\xi)|^2\,d\xi < \infty. \end{aligned}$$

Thus $\varphi_a(y(\,\cdot\,))(\,\cdot\,) \in L_2(\mathbb{R}^d)$ by Plancherel's theorem.

Let φ_a be such a map. We estimate the square of the functional in (14) to be maximized with $\varphi = \varphi_a$. To do this, we pass to Fourier images by Plancherel's theorem ($f(\xi) \leqslant \lambda(A)$ when $\xi \in \mathbb{R}^d \setminus A$):

$$
\begin{aligned}
&\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\bigl|(i\xi)^\alpha Fx(\xi)-\varphi_a(Fx(\,\cdot\,)|_A)(\xi)\bigr|^2\,d\xi\\
&\quad=\frac{1}{(2\pi)^d}\int_A\bigl|(i\xi)^\alpha Fx(\xi)-(i\xi)^\alpha a(\xi)Fx(\xi)\bigr|^2\,d\xi+\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d\setminus A}|\xi|^{2\alpha}\,|Fx(\xi)|^2\,d\xi\\
&\quad=\frac{1}{(2\pi)^d}\int_A|1-a(\xi)|^2\frac{|\xi|^{2\alpha}}{(1+\|\xi\|^2)^r}(1+\|\xi\|^2)^r\,|Fx(\xi)|^2\,d\xi\\
&\qquad\qquad+\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d\setminus A}\frac{|\xi|^{2\alpha}}{(1+\|\xi\|^2)^r}(1+\|\xi\|^2)^r\,|Fx(\xi)|^2\,d\xi\\
&\quad\leqslant \operatorname*{vrai\,sup}_{\xi\in A}|1-a(\xi)|^2 f(\xi)\frac{1}{(2\pi)^d}\int_A(1+\|\xi\|^2)^r\,|Fx(\xi)|^2\,d\xi\\
&\qquad\qquad+\lambda(A)\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d\setminus A}(1+\|\xi\|^2)^r\,|Fx(\xi)|^2\,d\xi. \qquad (15)
\end{aligned}
$$

It follows that if

$$
|1-a(\xi)|^2 f(\xi) \leqslant \lambda(A) \qquad (16)
$$

for a. e. $\xi \in \mathbb{R}^d$, then the right-hand side of (15) does not exceed $\lambda(A)$ and, therefore, the error in the method φ_a does not exceed $\sqrt{\lambda(A)}$. Using (10), we then have

$$
\sqrt{\lambda(A)} \leqslant E\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A, 0\bigr) \leqslant e\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A, 0, \varphi_a\bigr) \leqslant \sqrt{\lambda(A)},
$$

that is, $E(D^\alpha, H_2^r(\mathbb{R}^d), A, 0) = \sqrt{\lambda(A)}$, and φ_a is an optimal method.

The existence of such functions $a(\,\cdot\,)$ is obvious. For example, one can take the function which is identically equal to unity.

If $\lambda(A) = \widehat{\lambda}$, then $f(\xi) \leqslant \widehat{\lambda}$ for all $\xi \in \mathbb{R}^d$ and (16) holds with $a(\,\cdot\,) = 0$. Hence the zero method is optimal in this case.

When the function $\xi \mapsto (i\xi)^\alpha a(\xi) Fx(\xi)$ belongs to $L_1(A)$ (for example, when the measure of A is finite), the expression for the optimal method in the theorem is just the Fourier inversion formula. Otherwise the integral should be understood as the principal value for every $t \in \mathbb{R}^d$.

Case (b): $\delta > 0$, $\lambda(A) = \widehat{\lambda}$. We claim that the zero method is optimal in this situation.

Indeed,

$$
\begin{aligned}
&E\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A, \delta\bigr) \leqslant e\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A, \delta, 0\bigr)\\
&\quad= \sup_{\substack{x(\,\cdot\,)\in H_2^r(\mathbb{R}^d),\, y(\,\cdot\,)\in \widehat{L}_2(A)\\ \|Fx(\,\cdot\,)-y(\,\cdot\,)\|_{\widehat{L}_2(A)}\leqslant\delta}} \|D^\alpha x(\,\cdot\,)\|_{L_2(\mathbb{R}^d)} \leqslant \sup_{x(\,\cdot\,)\in H_2^r(\mathbb{R}^d)} \|D^\alpha x(\,\cdot\,)\|_{L_2(\mathbb{R}^d)}. \qquad (17)
\end{aligned}
$$

By Plancherel's theorem, the squared right-hand side of (17) is equal to the value of the problem

$$
\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}|\xi|^{2\alpha}\,|Fx(\xi)|^2\,d\xi \to \max, \quad \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}(1+\|\xi\|^2)^r\,|Fx(\xi)|^2\,d\xi \leqslant 1. \qquad (18)
$$

Hence, by the definition of $\widehat{\lambda}$, we have

$$\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}|\xi|^{2\alpha}\,|Fx(\xi)|^2\,d\xi=\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\frac{|\xi|^{2\alpha}}{(1+\|\xi\|^2)^r}(1+\|\xi\|^2)^r\,|Fx(\xi)|^2\,d\xi\leqslant\widehat{\lambda},$$

that is, the value of the problem (18) does not exceed $\widehat{\lambda}$. Therefore, in the cases $\delta\geqslant\widehat{\delta}$ and $0\leqslant\delta<\widehat{\delta}$, $\lambda(A)=\widehat{\lambda}$ (see (7) and (9)) we have

$$\sqrt{\widehat{\lambda}}\leqslant E\bigl(D^\alpha,H_2^r(\mathbb{R}^d),A,\delta\bigr)\leqslant e\bigl(D^\alpha,H_2^r(\mathbb{R}^d),A,\delta,0\bigr)\leqslant\sqrt{\widehat{\lambda}},$$

that is, $E(D^\alpha,H_2^r(\mathbb{R}^d),A,\delta)=\sqrt{\widehat{\lambda}}$ and $\widehat{\varphi}=0$ is an optimal method.

Case (c): $0<\delta<\widehat{\delta}$, $0\leqslant\lambda(A)<\widehat{\lambda}$. In this situation, our lower bounds for the optimal recovery error (see (8) and (12)) can be combined into one formula:

$$E\bigl(D^\alpha,H_2^r(\mathbb{R}^d),A,\delta\bigr)\geqslant\sqrt{\lambda_1\delta^2+\lambda_2},\tag{19}$$

where $\lambda_i>0$, $i=1,2$.

Let $a(\,\cdot\,)$ be a measurable function on A such that the function $\xi\mapsto(i\xi)^\alpha a(\xi)$ belongs to $L_\infty(A)$. As above, we search for optimal methods among those maps $\varphi_a\colon \operatorname{Im}I^\delta(A)\to L_2(\mathbb{R}^d)$ which act on the Fourier images by the rule $F\varphi_a(y(\,\cdot\,))(\xi)=(i\xi)^\alpha\widetilde{a}(\xi)\widetilde{y}(\xi)$ for a. e. $\xi\in\mathbb{R}^d$, where $\widetilde{a}(\,\cdot\,)=a(\,\cdot\,)$, $\widetilde{y}(\,\cdot\,)=y(\,\cdot\,)$ on A and $\widetilde{a}(\,\cdot\,)=0$, $\widetilde{y}(\,\cdot\,)=0$ outside A. Clearly, $\varphi_a(\,\cdot\,)\in L_2(\mathbb{R}^d)$.

Let φ_a be such a method. We estimate the value of the problem (13) in this situation. Passing by Plancherel's theorem to Fourier images, we see that the square of this value is equal to the value of the following problem:

$$\frac{1}{(2\pi)^d}\int_A\bigl|(i\xi)^\alpha Fx(\xi)-(i\xi)^\alpha a(\xi)y(\xi)\bigr|^2\,d\xi+\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d\setminus A}|\xi|^{2\alpha}\,|Fx(\xi)|^2\,d\xi\to\max,$$
$$\frac{1}{(2\pi)^d}\int_A|Fx(\xi)-y(\xi)|^2\,d\xi\leqslant\delta^2,\tag{20}$$
$$\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}(1+\|\xi\|^2)^r\,|Fx(\xi)|^2\,d\xi\leqslant1,\qquad x(\,\cdot\,)\in\mathcal{H}_2^r(\mathbb{R}^d),\quad y(\,\cdot\,)\in\widehat{L}_2(A).$$

We now estimate the first integrand in this functional by the Cauchy–Bunyakovskii inequality for every $\xi\in A$:

$$\begin{aligned}\bigl|(i\xi)^\alpha Fx(\xi)-(i\xi)^\alpha a(\xi)y(\xi)\bigr|^2&=|\xi|^{2\alpha}\,\bigl|(1-a(\xi))Fx(\xi)+a(\xi)(Fx(\xi)-y(\xi))\bigr|^2\\&\leqslant|\xi|^{2\alpha}\biggl(\frac{|1-a(\xi)|^2}{\lambda_2(1+\|\xi\|^2)^r}+\frac{|a(\xi)|^2}{\lambda_1}\biggr)\bigl(\lambda_2(1+\|\xi\|^2)^r\,|Fx(\xi)|^2\\&\qquad+\lambda_1|Fx(\xi)-y(\xi)|^2\bigr).\end{aligned}\tag{21}$$

We put

$$S_a=\operatorname*{vrai\,sup}_{\xi\in A}|\xi|^{2\alpha}\biggl(\frac{|1-a(\xi)|^2}{\lambda_2(1+\|\xi\|^2)^r}+\frac{|\alpha(\xi)|^2}{\lambda_1}\biggr).$$

Assuming that $S_a \leqslant 1$, we can integrate the inequality (21) over A and take into account that $f(\xi) \leqslant \lambda(A)$ outside A in order to obtain the following estimate for the functional in (20):

$$\lambda_2 \frac{1}{(2\pi)^d} \int_A (1+\|\xi\|^2)^r\, |Fx(\xi)|^2\, d\xi + \lambda_1 \frac{1}{(2\pi)^d} \int_A |Fx(\xi) - y(\xi)|^2\, d\xi \\ + \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d\setminus A} \frac{|\xi|^{2\alpha}}{(1+\|\xi\|^2)^r} (1+\|\xi\|^2)^r\, |Fx(\xi)|^2\, d\xi \\ \leqslant \lambda_2 \frac{1}{(2\pi)^d} \int_A (1+\|\xi\|^2)^r\, |Fx(\xi)|^2\, d\xi + \lambda_1 \frac{1}{(2\pi)^d} \int_A |Fx(\xi) - y(\xi)|^2\, d\xi \\ + \lambda(A) \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d\setminus A} (1+\|\xi\|^2)^r\, |Fx(\xi)|^2\, d\xi. \qquad (22)$$

If $\lambda(A) \leqslant h(\delta)$, then $\Delta = \delta$ and, therefore, $\lambda(A) \leqslant h(\Delta) = \lambda_2$. But if $\lambda(A) \geqslant h(\delta)$, then $\lambda_2 = h(\Delta) = \lambda(A)$. Thus we always have $\lambda(A) \leqslant \lambda_2$. Therefore, using the constraints in the problem (20), we obtain from (22) that the functional to be maximized in (20) does not exceed $\lambda_2 + \lambda_1\delta^2$. Hence the error in the method φ_a does not exceed $\sqrt{\lambda_2 + \lambda_1\delta^2}$. Together with (19), this means that the method φ_a is optimal.

We now prove the existence of functions $a(\,\cdot\,)$ such that $S_a \leqslant 1$.

If the inequality

$$|\xi|^{2\alpha} \biggl(\frac{|1 - a(\xi)|^2}{\lambda_2 (1+\|\xi\|^2)^r} + \frac{|\alpha(\xi)|^2}{\lambda_1} \biggr) \leqslant 1 \qquad (23)$$

holds for a. e. $\xi \in \mathbb{R}^d$, then $S_a \leqslant 1$. Furthermore, if

$$-|\xi|^{2\alpha} + \lambda_1 + \lambda_2 (1+\|\xi\|^2)^r \geqslant 0 \qquad \forall\, \xi \in \mathbb{R}^d, \qquad (24)$$

then we easily see (by completing the square) that (23) is equivalent to the corresponding hypothesis of the theorem. It follows that for every measurable function $a(\,\cdot\,)$ satisfying the hypotheses of the theorem, the function $\xi \mapsto (i\xi)^\alpha a(\xi)$ belongs to $L_\infty(A)$, and hence there are 'sufficiently many' optimal methods.

It remains to prove (24). To do this, we define a function $g(\,\cdot\,)$ on the half-line $[0,\infty)$ by the formula

$$g(x) = \frac{\alpha^\alpha}{\overline{\alpha}^{\overline{\alpha}}} (x^{1/r} - 1)^{\overline{\alpha}}.$$

We easily see that this function is concave on $[x_0, \infty)$, where $x_0 = (r/(r - \overline{\alpha}))^r$.

In our case, $\Delta < \widehat{\delta}$ or, equivalently, $\Delta^{-2} > x_0$. A direct calculation shows that the line $x \mapsto \lambda_1 + \lambda_2 x$ is tangent to the graph of $g(\,\cdot\,)$ at the point Δ^{-2}. Since $g(\,\cdot\,)$ is concave, it follows that

$$g(x) \leqslant \lambda_1 + \lambda_2 x \qquad \forall\, x \geqslant x_0.$$

Suppose that $\xi \in \mathbb{R}^d$ and put $x_\xi = (1+\|\xi\|^2)^r$. Then

$$g(x_\xi) = \frac{\alpha^\alpha}{\overline{\alpha}^{\overline{\alpha}}} \|\xi\|^{2\overline{\alpha}}.$$

It follows from the inequality between the arithmetic and geometric means (see [19], Russian p. 29) that

$$|\xi|^{2\alpha} \leqslant \frac{\alpha^\alpha}{\overline{\alpha}^{\overline{\alpha}}}\|\xi\|^{2\overline{\alpha}}.$$

Combining the last two formulae and taking into account that $x_\varepsilon > x_0$, we obtain

$$|\xi|^{2\alpha} \leqslant \frac{\alpha^\alpha}{\overline{\alpha}^{\overline{\alpha}}}\|\xi\|^{2\overline{\alpha}} = g(x_\xi) \leqslant \lambda_1 + \lambda_2 x_\xi = \lambda_1 + \lambda_2(1+\|\xi\|^2)^r.$$

This proves (24) and hence establishes the expression for the functions $a(\,\cdot\,)$ in the statement of the theorem.

As in the case $\delta = 0$, when the function $\xi \mapsto (i\xi)^\alpha a(\xi) Fx(\xi)$ belongs to $L_1(A)$ (for example, when the measure of A is finite), the expression for the optimal method in the theorem is the Fourier inversion formula. Otherwise the integral should be understood as the principal value for every $t \in \mathbb{R}^d$. □

Proof of Theorem 2. Suppose that $A_\sigma \in \mathcal{A}_\sigma$. We claim that $\lambda(A_\sigma) \geqslant \lambda(\sigma)$. Indeed, if $\lambda(A_\sigma) < \lambda(\sigma)$, then there is a $\lambda < \lambda(\sigma)$ such that $\operatorname{mes}(A_\sigma \cap \Omega_\lambda) = \operatorname{mes}\Omega_\lambda$. Since $\operatorname{mes}\Omega_\lambda > \operatorname{mes}\Omega_{\lambda(\sigma)} = \sigma$, we have $\operatorname{mes} A_\sigma > \sigma$. This is impossible.

Suppose that $\delta = 0$. Then it follows from Theorem 1 that

$$E\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A_\sigma, 0\bigr) = \sqrt{\lambda(A_\sigma)} \geqslant \sqrt{\lambda(\sigma)}.$$

But since $\operatorname{mes}\Omega_{\lambda(\sigma)} = \sigma$ and $\lambda(\Omega_{\lambda(\sigma)}) = \lambda(\sigma)$, the set $\Omega_{\lambda(\sigma)}$ is optimal and, clearly, every set which differs from $\Omega_{\lambda(\sigma)}$ only by a set of measure zero is also optimal.

If $\delta \geqslant \widehat{\delta}$, then for every set (and, in particular, for every A_σ) we have

$$E\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A_\sigma, 0\bigr) = \sqrt{\widehat{\lambda}}.$$

Hence every set is optimal in this case.

We now suppose that $0 < \delta < \widehat{\delta}$ and $A_\sigma \in \mathcal{A}_\sigma$. If $\lambda(\sigma) \geqslant h(\delta)$, then $\lambda(\sigma,\delta) = \lambda(\sigma)$. It was proved above that $\lambda(A_\sigma) \geqslant \lambda(\sigma)$ and, therefore, $\lambda(A_\sigma) \geqslant h(\delta)$. In this case, as noted in comment 2 on Theorem 1, the optimal recovery error decreases as $\lambda(A)$ decreases. Since $\lambda(A_\sigma) \geqslant \lambda(\sigma) = \lambda(\Omega_{\lambda(\sigma)})$, we have

$$E\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A_\sigma, \delta\bigr) \geqslant E\bigl(D^\alpha, H_2^r(\mathbb{R}^d), \Omega_{\lambda(\sigma)}, \delta\bigr).$$

Therefore $\Omega_{\lambda(\sigma)}$ is an optimal set.

Suppose that $\lambda(\sigma) \leqslant h(\delta)$. Then $\lambda(\sigma,\delta) = h(\delta)$. If $\lambda(A_\sigma) \geqslant h(\delta)$, then the argument used in the previous case shows that

$$E\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A_\sigma, \delta\bigr) \geqslant E\bigl(D^\alpha, H_2^r(\mathbb{R}^d), \Omega_{h(\delta)}, \delta\bigr).$$

But if $\lambda(A_\sigma) \leqslant h(\delta)$, then, by Theorem 1,

$$E\bigl(D^\alpha, H_2^r(\mathbb{R}^d), A_\sigma, \delta\bigr) = \sqrt{\biggl(\frac{r}{\overline{\alpha}}(\widehat{\delta}^{2/r} - \delta^{2/r}) + 1\biggr) h(\delta)} = E\bigl(D^\alpha, H_2^r(\mathbb{R}^d), \Omega_{h(\delta)}, \delta\bigr),$$

whence $\Omega_{h(\delta)}$ is an optimal set.

Thus, for $\delta > 0$, the set $\Omega_{\lambda(\sigma,\delta)}$ is optimal, and so is every set which differs from $\Omega_{\lambda(\sigma,\delta)}$ by a set of measure zero. □

We now consider an example. Take $\alpha = (1, 0, \dots, 0)$ and $r = 2$. In other words, consider the problem of recovering the partial derivative $x_{t_1}(\,\cdot\,)$ on the class $H_2^2(\mathbb{R}^d)$. In this case,

$$\widehat{\lambda} = \frac{1}{4}, \qquad \widehat{\delta} = \frac{1}{2},$$

$$\Omega_\lambda = \left\{ (\xi_1, \dots, \xi_d) \in \mathbb{R}^d \;\middle|\; \frac{\xi_1^2}{(1 + \xi_1^2 + \dots + \xi_d^2)^2} \geqslant \lambda \right\}.$$

We easily see that the set Ω_λ with $\lambda < 1/4$ consists of two balls:

$$\Omega_\lambda = \left\{ (\xi_1, \dots, \xi_d) \in \mathbb{R}^d \;\middle|\; \left(|\xi_1| - \frac{1}{2\sqrt{\lambda}} \right)^2 + \xi_2^2 + \dots + \xi_d^2 \leqslant \frac{1}{4\lambda} - 1 \right\}.$$

By Theorem 1, $E(D^\alpha, H_2^2(\mathbb{R}^d), A, \delta) = 1/2$ when $\delta \geqslant 1/2$. If $\delta < 1/2$, then

$$E\big(D^\alpha, H_2^2(\mathbb{R}^d), A, \delta\big) = \begin{cases} \sqrt{(1-\delta)\delta}, & \lambda(A) \leqslant \dfrac{\delta}{2}, \\ \sqrt{\left(\dfrac{1 - 4\lambda(A)}{4\lambda(A)} \right) \delta^2 + \lambda(A)}, & \lambda(A) > \dfrac{\delta}{2}. \end{cases}$$

A family of optimal methods in this case can easily be obtained from Theorem 1. In particular, the method

$$\widehat{\varphi}(Fx(\,\cdot\,)|_A)(t) = \frac{1}{(2\pi)^d} \int_A \frac{i\xi_1}{1 + \frac{\Delta^2}{1-2\Delta}(1 + \xi_1^2 + \dots + \xi_d^2)^2} Fx(\xi) e^{\langle \xi, t \rangle}\, d\xi$$

is optimal for $0 < \delta < 1/2$, where

$$\Delta = \begin{cases} \delta, & \lambda(A) \leqslant \dfrac{\delta}{2}, \\ 2\lambda(A), & \lambda(A) > \dfrac{\delta}{2}. \end{cases}$$

By the well-known formula for the volume of a d-dimensional ball we have

$$\operatorname{mes} \Omega_\lambda = \frac{2\pi^{d/2}}{\Gamma(d/2 + 1)} \left(\frac{1}{4\lambda} - 1 \right)^{d/2}$$

and, therefore,

$$\lambda(\sigma) = \frac{1}{4} \left(1 + \frac{1}{\pi} \left(\frac{\sigma}{2} \Gamma\left(\frac{d}{2} + 1 \right) \right)^{2/d} \right)^{-1}.$$

It was mentioned in the proof of Theorem 2 that all sets are optimal for $\delta \geqslant 1/2$. When $\delta < 1/2$, the same theorem shows that the balls

$$\left\{ (\xi_1, \dots, \xi_d) \in \mathbb{R}^d \;\middle|\; \left(|\xi_1| - \sqrt{R^2 + 1} \right)^2 + \xi_2^2 + \dots + \xi_d^2 \leqslant R^2 \right\}$$

are optimal sets, where

$$R=\begin{cases}\dfrac{1}{\sqrt{\pi}}\left(\dfrac{\sigma}{2}\Gamma\left(\dfrac{d}{2}+1\right)\right)^{1/d}, & \sigma<\dfrac{2}{\pi^{d/2}\Gamma(d/2+1)}\left(\dfrac{1}{2\delta}-1\right)^{d/2},\\ \sqrt{\dfrac{1}{2\delta}-1}, & \sigma\geqslant\dfrac{2}{\pi^{d/2}\Gamma(d/2+1)}\left(\dfrac{1}{2\delta}-1\right)^{d/2}.\end{cases}$$

Bibliography

[1] A. Kolmogoroff, "Über die beste Annäherung von Functionen einer gegebenen Functionenklasse", *Ann. of Math.* (2) **37**:1 (1936), 107–110.

[2] A. Sard, "Best approximate integration formulas; best approximation formulas", *Amer. J. Math.* **71**:1 (1949), 80–91.

[3] S. M. Nikol'skii, "On the question of estimates for approximation by quadrature formulae", *Uspekhi Mat. Nauk* **5**:2(36) (1950), 165–177. (Russian)

[4] S. A. Smolyak, *On optimal recovery of functions and functionals of them*, Diss. . . . Cand. Phys.-Math. Sci., Moscow State Univ., Moscow 1965. (Russian)

[5] C. A. Micchelli and T. J. Rivlin, "A survey of optimal recovery", *Optimal estimation in approximation theory*, Proc. Internat. Sympos. (Freudenstadt 1976), Plenum, New York 1977, pp. 1–54.

[6] J. F. Traub and H. Woźniakowski, *A general theory of optimal algorithms*, ACM Monograph Series, Academic Press, New York–London 1980; Russian transl., Mir, Moscow 1983.

[7] C. A. Micchelli and T. J. Rivlin, "Lectures on optimal recovery", *Numerical analysis, Lancaster* 1984 (Lancaster 1984), Lecture Notes in Math., vol. 1129, Springer-Verlag, Berlin 1984, pp. 21–93.

[8] V. V. Arestov, "Optimal recovery of operators and related problems", *Proc. All-Union School in Function Theory* (Dushanbe, August 1986), Trudy Mat. Inst. Steklov., vol. 189, Nauka, Moscow 1989, pp. 3–20; English transl., *Proc. Steklov Inst. Math.* **189**:4 (1990), 1–20.

[9] K. Yu. Osipenko, *Optimal recovery of analytic functions*, Nova Science Publ., Inc., Huntington, New York 2000.

[10] G. G. Magaril-Il'yaev and V. M. Tikhomirov, *Convex analysis: theory and applications*, Editorial URSS, Moscow 2011; English transl. of 1st ed., Transl. Math. Monogr., vol. 222, Amer. Math. Soc., Providence, RI 2003.

[11] G. G. Magaril-Il'yaev and V. M. Tikhomirov, *Convex analysis: theory and applications*, Transl. Math. Monogr., vol. 222, Amer. Math. Soc., Providence, RI 2003.

[12] G. G. Magaril-Il'yaev and K. Yu. Osipenko, "Optimal recovery of functions and their derivatives from inaccurate information about the spectrum and inequalities for derivatives", *Funktsional. Anal. i Prilozhen.* **37**:3 (2003), 51–64; English transl., *Funct. Anal. Appl.* **37**:3 (2003), 203–214.

[13] G. G. Magaril-Il'yaev and K. Yu. Osipenko, "On the reconstruction of convolution-type operators from inaccurate information", *Function Theory and Differential Equations*, Collection of papers, Tr. Mat. Inst. Steklova, vol. 269, MAIK, Moscow 2010, pp. 181–192; English transl., *Proc. Steklov Inst. Math.* **269** (2010), 174–185.

[14] G. G. Magaril-Il'yaev and K. Yu. Osipenko, "On optimal harmonic synthesis from inaccurate spectral data", *Funktsional. Anal. i Prilozhen.* **44**:3 (2010), 76–79; English transl., *Funct. Anal. Appl.* **44**:3 (2010), 223–225.

[15] G. G. Magaril-Il'yaev and K. Yu. Osipenko, "On the best choice of information in the problem of recovering a function from its spectrum", *Math. Forum*, vol. 1: *Investigations in Math. Analysis*, VSC RAS, Vladikavkaz 2008, pp. 142–150. (Russian)

[16] G. G. Magaril-Il'yaev and K. Yu. Osipenko, "How best to recover a function from its inaccurately given spectrum?", *Mat. Zametki* **92**:1 (2012), 59–67; English transl., *Math. Notes* **92**:1 (2012), 51–58.

[17] G. G. Magaril-Il'yaev and E. O. Sivkova, "Best recovery of the Laplace operator of a function from incomplete spectral data", *Mat. Sb.* **203**:4 (2012), 119–130; English transl., *Sb. Math.* **203**:4 (2012), 569–580.

[18] G. G. Magaril-Il'yaev and K. Yu. Osipenko, "Optimal recovery of derivatives on Sobolev classes", *Vladikavk. Mat. Zh.* **5**:1 (2003), 39–47. (Russian)

[19] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Univ. Press, Cambridge 1934; Russian transl., IL, Moscow 1948.

Georgii G. Magaril-Il'yaev
Moscow State University
Institute of Information Transmission Problems
(Kharkevich Institute), RAS, Moscow
E-mail: magaril@mech.math.msu.su

Received 25/OCT/13
Translated by A. V. DOMRIN

Konstantin Yu. Osipenko
Moscow State University
Russian State University of Technology (MATI), Moscow
Institute of Information Transmission Problems
(Kharkevich Institute), RAS, Moscow
E-mail: kosipenko@yahoo.com