Statistical approach to inverse boundary problems for partial differential equations

Golubev G. and Khasminskii R.[1]

Weierstraß–Institut für Angewandte Analysis and Stochastik
Mohrenstaße 39
D–10117, Berlin, Deutchland
e-mail: golubev@wias-berlin.de

Wayne State University
Detroit, MI 48202, USA
e-mail: rafail@math.wayne.edu

[1]This work was partially supported NSF grant DMS 9600245 and ONR grant N00014-95-0793
1991 Mathematics Subject Classifications. 62G05.
Key words and phrases. Partial differential equation, minimax estimator, boundary problem.

Abstract

Inverse problems for elliptic and parabolic partial differential equations are considered. It is assumed that a solution of the equation is observed in white Gaussian noise with a small spectral density. The goal is to recover smooth but unknown boundary or initial conditions based on the noisy data. It is shown that the second order minimax estimators are linear as the spectral density of the noise goes to zero.

1 Introduction

Ill-posed inverse problems for ordinary and partial differential equations are very popular in modern mathematics. We refer to the books Hadamard (1932), Ivanov & Vasin & Tanana (1978), Tihonov & Arsenin (1978), Lavrentiev & Romanov & Shishatskii (1980). Usually these problems are associated with recovering of an unknown differential operator or unknown boundary conditions based on the observations of a solution of the equation in some domain. There are two important questions, which arise naturally in this context. In the first place one wants to prove the uniqueness of a solution of the inverse boundary problem if it is known that the restored boundary functions belong to a certain functional class. If the solution represents a noisy data, one has to show next that the error in the recovered functions tends to zero when the noise in the data tends to zero too, Lattes & Lions (1967).

Statistical approach to inverse problems is based on the assumption that solutions are observed in a white noise with a small spectral density $\varepsilon \to 0$. The goal is to find estimators which are at least consistent. The main advantage of the statistical approach lies in the fact that it helps to compare different approaches to inverse problems. Thus we have a mathematical basis for choosing solutions, which are optimal in different senses. Sudakov & Khalfin (1964) and Khalfin (1978) were among the first who propose to use the statistical approach to ill-posed problems. In Chow & Khasminskii (1997a) and Chow & Ibragimov & Khasminskii (1997b) this approach was used for estimation of a source based on solutions of ordinary and partial differential equations.

In the present paper we consider from statistical point of view two very popular inverse problems:

i) recovering of initial conditions for parabolic equations based on the observations in a fixed time strip,

ii) recovering of boundary conditions for elliptic equations based on the observations in an internal domain.

Many scientists have paid attention on these problems. For instance, the major part of the monograph Lattes & Lions (1967) is devoted to the problem i). From statistical point of view the authors have proposed very interesting estimates and demonstrated that they are consistent.

We show in this paper how to construct estimates with optimal rates of convergence when the spectral density of the noise goes to zero ($\varepsilon \to 0$). Moreover we indicate these rates up to constants. Here we follow the well-known paper Pinsker (1980). Unfortunately these rates are very low and there exist many estimators having the optimal rates. In order to discriminate between the estimators in this situation we find the second order term in the expansion of the minimax risk and propose second order minimax estimators. In the problem of estimation initial conditions for parabolic equations we propose also exponentially minimax estimators.

1.1 Estimation of initial conditions for parabolic equations

To simplify technical details we consider the problem in the simplest setting. Let $u(t,x)$ be a solution of the heat conductivity equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \tag{1}$$

in the domain $\mathcal{D} = \{0 \le x \le 1\} \times \{t > 0\}$ with the periodic boundary conditions

$$u(t,0) = u(t,1), \quad \frac{\partial u}{\partial x}(t,0) = \frac{\partial u}{\partial x}(t,1), \tag{2}$$

and with the initial condition

$$u(0,x) = \theta(x). \tag{3}$$

Identifying points $x = 0$ and $x = 1$, one could say that the equation (1) is considered on the surface of the cylinder $\mathcal{C} \times \{t > 0\}$, where $\mathcal{C}$ is the circle of the unit length. It is also assumed that $\theta(x),\ x \in \mathcal{C}$ has Sobolev smoothness β, so

$$\theta(x) = \sum_{j=-\infty}^{\infty} \theta_j \exp(2\pi \mathrm{i} jx), \quad \mathrm{i} = \sqrt{-1}, \tag{4}$$

where $\theta_j = \theta^*_{-j}$ and

$$\sum_{j=-\infty}^{\infty} |\theta_j|^2 |2\pi j|^{2\beta} \le L. \tag{5}$$

The set of all function satisfying to (4) and (5) constitutes the Sobolev class $\mathcal{W}_2^\beta(L)$. Next we consider two models of observations.

*Model 1^a^*. Assume that a solution of the equation (1) is observed at $t = T$ in a white Gaussian noise with the spectral density ε^2. Thus we observe the generalized random process $Y(x),\ x \in [0,1]$, which has the form

$$Y(x) = u(T,x) + \varepsilon n(x), \tag{6}$$

where $n(x)$ is a white Gaussian noise, t.e. a generalized Gaussian process with the covariance function $\mathbf{E} n(x)n(y) = \delta(x-y)$; here $\delta(\cdot)$ is the Dirac δ-function. It means that for any $\varphi_\alpha(\cdot) \in L_2(\mathcal{C})$ we can observe the family of random variables

$$Y_\alpha = \int_{\mathcal{C}} u(T,x)\varphi_\alpha(x)\,\mathrm{d}x + \varepsilon \xi_\alpha, \tag{7}$$

where ξ_α are Gaussian random variables with zero mean and the covariance function

$$\mathbf{E}\xi_{\alpha_1}\xi_{\alpha_2} = \int_{\mathcal{C}} \varphi_{\alpha_1}(x)\varphi_{\alpha_2}(x)\,\mathrm{d}x.$$

Thus, our problem is to find the minimax risk

$$r_\varepsilon(\mathcal{W}_2^\beta(L)) = \inf_{\tilde{\theta}} \sup_{\theta\in W_2^\beta(L)} \mathbf{E}_\theta \int_{\mathcal{C}} \left(\tilde{\theta}(x,Y) - \theta(x)\right)^2 \mathrm{d}x,$$

where the *inf* is taken over all estimators based on observations (6). We want also to find a minimax estimator i.e. the estimator θ^* such that

$$r_\varepsilon(\mathcal{W}_2^\beta(L)) = \sup_{\theta\in W_2^\beta(L)} \mathbf{E}_\theta \int_{\mathcal{C}} \left(\theta^*(x,Y) - \theta(x)\right)^2 \mathrm{d}x.$$

The above problem can be easily reduced to the following one. Let $u(t,x)$ be a solution of (1)–(3). Since this function is periodic, it can be expanded into the Fourier series

$$u(t,x) = \sum_{j=-\infty}^{\infty} u_j(t)\exp(2\pi \mathrm{i} jx). \tag{8}$$

Using the Fourier method (cf. Petrovskii (1950), s. 38) one obtains

$$u_j(t) = \theta_j \exp(-4\pi^2 j^2 t), \tag{9}$$

where θ_j are the Fourier coefficients associated with the function $\theta(x)$ (see (4)).

Since the Fourier basis is a complete orthonormal system in $L_2(\mathcal{C})$, the problem of estimation $\theta(x)$ is equivalent to estimation of the Fourier coefficients $\theta_j,\ j = 0, \pm 1, \pm 2, \ldots$ based on the observations

$$Y_j = \theta_j \exp(-4\pi^2 j^2 T) + \varepsilon\xi_j, \tag{10}$$

where ξ_j are independent $\mathcal{N}(0,1)$. Or equivalently, we have to estimate θ_j based on the data

$$Z_j = \theta_j + \varepsilon\sigma_j\xi_j, \tag{11}$$

where $\sigma_j = \exp(4\pi^2 j^2 T)$ and θ_j belong to the following set

$$\Theta = \left\{\theta_j :\ \sum_{j=-\infty}^{\infty} |\theta_j|^2 |2\pi j|^{2\beta} \le L\right\}. \tag{12}$$

Note the minimax risk admits the following representation in the terms of the Fourier coefficients

$$r_\varepsilon(\mathcal{W}_2^\beta(L)) = \inf_{\tilde{\theta}} \sup_{\theta\in\Theta} \mathbf{E}_\theta \sum_{j=-\infty}^{\infty} \left(\tilde{\theta}_j - \theta_j\right)^2. \tag{13}$$

Model 1^b. It is assumed that the solution of (1) is observed into the strip $T \le t \le T+h$ in a two-dimensional white Gaussian noise

$$Y(t,x) = u(t,x) + \varepsilon n(t,x), \quad t\in[T,T+h],\ x\in\mathcal{C},$$

where $n(t,x)$ is a generalized two-dimensional Gaussian field with the covariance function

$$\mathbf{E}n(t_1,x_1)n(t_2,x_2) = \delta(t_1-t_2, x_1-x_2).$$

By (8) and (9) we get the equivalent model (see also (7))

$$\tilde{Z}_j = \theta_j + \varepsilon\tilde{\sigma}_j\xi_j, \tag{14}$$

where

$$\begin{aligned}\tilde{\sigma}_j &= \left(\int_T^{T+h} \exp(-8\pi^2 j^2 t)\,\mathrm{d}t\right)^{-1/2} \\ &= 2\sqrt{2}\pi|j|\exp(4\pi^2 j^2 T)\left(1-\exp(-8\pi^2 j^2 h)\right)^{-1/2}.\end{aligned}$$

Based on the observation from (14) we have to find the minimax estimator and to calculate its risk (cf. (13)). Thus we see that the only difference between models 1^a and 1^b is in the definition of σ_j^2 and $\tilde{\sigma}_j^2$.

1.2 Estimation of boundary conditions for elliptic equations

Consider the simplest case of the Dirichlet problem for the Laplace equation on the circle of radius 1

$$\begin{aligned}&\Delta u = 0 \\ &u(\cos\varphi, \sin\varphi) = f(\varphi).\end{aligned} \tag{15}$$

It is known (see e.g. Petrovskii (1950), s. 29) that the solution of the above problem can be rewritten in the polar coordinates as

$$u(r,\varphi) = \frac{\theta_0}{\sqrt{2\pi}} + \frac{1}{\sqrt{\pi}}\sum_{k=1}^{\infty} r^k(\theta_k\cos(k\varphi) + \theta_{-k}\sin(k\varphi)), \tag{16}$$

where θ_k are the Fourier coefficients of $f(\varphi)$, so that

$$f(\varphi) = \frac{\theta_0}{\sqrt{2\pi}} + \frac{1}{\sqrt{\pi}}\sum_{k=1}^{\infty}(\theta_k\cos(k\varphi) + \theta_{-k}\sin(k\varphi)).$$

Let us assume that a solution of (15) is observed on the circle $\mathcal{C}_\rho$ of the radius $\rho < 1$ in a white Gaussian noise. Thus we observe a generalized Gaussian field

$$Y(r,\varphi) = ru(r,\varphi) + \varepsilon\sqrt{r}n(r,\varphi),\ 0 \le r \le \rho,\ \varphi \in [0,2\pi], \tag{17}$$

where $n(r,\varphi)$ is a white Gaussian noise i.e. a generalized Gaussian field with the covariance function

$$\mathbf{E}n(r_1,\varphi_1)n(r_2,\varphi_2) = \delta(r_1-r_2, \varphi_1-\varphi_2).$$

Substituting (16) in (17), multiplying it by 1, $\cos(k\varphi)$, $\sin(k\varphi)$ and integrating by parts over φ, we arrive to the equivalent family of the observations (17)

$$\theta_k(r) \quad = \quad \theta_k r^{|k|+1/2} + \varepsilon n_k(r), \;\; k = 0, \pm 1, \pm 2, ... \tag{18}$$

where $0 \le r \le \rho$ and $n_k(r)$ are one-dimensional independent (for different k) white Gaussian noises. Thus the problem of estimation θ_k based on (18) is equivalent to estimation of θ_k based on the observations

$$Y_k \quad = \quad \frac{\displaystyle\int_0^\rho r^{|k|+1/2}\theta_k(r)\,\mathrm{d}r}{\displaystyle\int_0^\rho r^{2|k|+1}\,\mathrm{d}r} = \theta_k + \varepsilon\left(\int_0^\rho r^{2|k|+1}\,\mathrm{d}r\right)^{-1/2}\xi_k,$$

where ξ_k are independent $\mathcal{N}(0,1)$. Finally we arrive to the following equivalent problem: to estimate θ_k based on the data

$$Y_k = \theta_k + \varepsilon\sigma_k\xi_k, \;\; k = 0, \pm 1, \pm 2, ...$$

where ξ_k are i.i.d. $\mathcal{N}(0,1)$ and

$$\sigma_k = \sqrt{2|k|+2}\rho^{-|k|-1}.$$

We see that each problem under consideration is equivalent to the problem of estimation unknown parameters θ_k based on the observations (11), while the prior information is provided by (12). The type of the considered problem is reflected only in the definition of σ_k. For parabolic equations σ_k grow like $\exp(Ck^2)$. In the elliptic case we have a growth of the order $\exp(Ck)$. This difference plays an essential role only when we prove that second order minimax estimators are linear for the both models.

2 Linear estimation

2.1 An upper bound for the minimax risk

Consider the following linear estimator

$$\hat{\theta}_j = h_j Y_j$$

of the parameters θ_j based on the observations (11). It is assumed θ_j are subjected to restrictions (12). The problem of calculation of the linear minimax risk

$$r_\varepsilon^{\mathcal{L}}(\mathcal{W}_2^\beta(L)) = \inf_{h_j}\sup_{\theta\in\Theta}\mathbf{E}_\theta\sum_{j=-\infty}^{\infty}(h_jY_j - \theta_j)^2$$

has the well-known solution, Pinsker (1980).

Proposition 1

$$r_\varepsilon^{\mathcal{L}}(\mathcal{W}_2^\beta(L)) = \varepsilon^2 \sum_{j=-\infty}^{\infty} \sigma_j^2 \left[1 - \left|\frac{j}{W}\right|^\beta\right]_+, \tag{19}$$

where W is a root of the equation

$$\varepsilon^2 \sum_{j=-\infty}^{\infty} \sigma_j^2 |2\pi j|^{2\beta} \left[\left|\frac{W}{j}\right|^\beta - 1\right]_+ = L. \tag{20}$$

The estimator

$$\theta_j^* = \left[1 - \left|\frac{j}{W}\right|^\beta\right]_+ Y_j \tag{21}$$

is the minimax linear estimator.

Proof follows directly from the fact that the functional

$$L[h,\theta] = \sum_{j=-\infty}^{\infty} (1-h_j)^2\theta_j^2 + \varepsilon^2 \sum_{j=-\infty}^{\infty} \sigma_j^2 h_j^2 = \mathbf{E}_\theta \sum_{j=-\infty}^{\infty} (h_j Y_j - \theta_j)^2$$

has a saddle point on $l_2(-\infty,\infty)\times\Theta$ (see e.g. Pinsker (1980)). The components of this saddle point are given by

$$h_j^0 = \left[1 - \left|\frac{j}{W}\right|^\beta\right]_+, \tag{22}$$

$$\left(\theta_j^0\right)^2 = \varepsilon^2\sigma_j^2 \left[\left|\frac{W}{j}\right|^\beta - 1\right]_+, \tag{23}$$

where $[x]_+ = \max(0,x)$.

2.2 A lower bound for the minimax risk

To show that the linear estimator defined by (20), (21) is a first order minimax estimator we will use the following fact.

Proposition 2 *Let $(\theta_j^0)^2$ are defined by (23). Then*

$$r_\varepsilon(\mathcal{W}_2^\beta(L)) \geq \sum_{j\neq 0}(\theta_j^0)^2 - \frac{4}{\varepsilon^2}\sum_{j\neq 0}(\theta_j^0)^4\sigma_j^{-2}. \tag{24}$$

Proof. Let θ_j be independent random variables taking values θ_j^0, $-\theta_j^0$ with the probabilities $1/2$. It is clear that

$$r_\varepsilon(\mathcal{W}_2^\beta(L)) \geq \sum_{j\neq 0} \inf_{\tilde\theta_j} \mathbf{E}\,\mathbf{E}\{(\tilde\theta_j - \theta_j)^2|\theta_j\}. \tag{25}$$

It is also evident that the *inf* in the right-hand side of (25) is attained when $\tilde\theta_j$ is the Bayesian estimator

$$\tilde\theta_j = \theta_j^0 \operatorname{th}\left(\frac{Y_j\theta_j^0}{\varepsilon^2\sigma_j^2}\right),$$

where

$$\operatorname{th} x = \frac{\exp(x)-\exp(-x)}{\exp(x)+\exp(-x)}.$$

Then we have

$$\begin{aligned}
\mathbf{E}\,\mathbf{E}\{(\tilde\theta_j-\theta_j)^2|\theta_j\} &= \mathbf{E}\{(\tilde\theta_j-\theta_j)^2|\theta_j=\theta_j^0\} \qquad (26)\\
&= (\theta_j^0)^2\,\mathbf{E}\left\{\left(1+\frac{1}{2}\left(\exp\left(\frac{2Y_j\theta_j^0}{\varepsilon^2\sigma_j^2}\right)-1\right)\right)^{-2}\middle|\,\theta_j=\theta_j^0\right\}\\
&\geq (\theta_j^0)^2\,\mathbf{E}\left\{\left(1-\frac{1}{2}\left(\exp\left(\frac{2Y_j\theta_j^0}{\varepsilon^2\sigma_j^2}\right)-1\right)\right)^{2}\middle|\,\theta_j=\theta_j^0\right\}\\
&= (\theta_j^0)^2\left(\frac{9}{4}-\frac{3}{2}\exp\left(\frac{4(\theta_j^0)^2}{\varepsilon^2\sigma_j^2}\right)+\frac{1}{4}\exp\left(\frac{12(\theta_j^0)^2}{\varepsilon^2\sigma_j^2}\right)\right)\\
&\geq (\theta_j^0)^2\left(\frac{5}{2}-\frac{3}{2}\exp\left(\frac{4(\theta_j^0)^2}{\varepsilon^2\sigma_j^2}\right)\right).
\end{aligned}$$

Since $5-3\exp(x)$ is a convex function, one easily obtains that

$$5-3\exp(x)\geq 2-x/\log(5/3) \qquad (27)$$

for $x\in[0,\log(5/3)]$. Note also that if

$$\frac{4(\theta_j^0)^2}{\varepsilon^2\sigma_j^2} > \log(5/3)$$

the right-hand side in (26) is negative. Then (27) and (26) yield

$$\mathbf{E}\,\mathbf{E}\{(\tilde\theta_j-\theta_j)^2|\theta_j\}\geq(\theta_j^0)^2\left(1-\frac{4(\theta_j^0)^2}{\varepsilon^2\sigma_j^2}\right), \qquad (28)$$

thus, by virtue of (25), proving the required inequality (24).

2.3 Asymptotic behaviour of the minimax risk

At the first glance it seems that the lower bound given by Proposition 2 is not very good. This observation is true in the case when σ_j grow not very rapidly. But in our models we have at least an exponential growth of σ_j, and this fact plays an essential role in the proving of the following theorem.

Theorem 1 *Let*

$$\frac{|\sigma_{j+1}|}{|\sigma_j|} \geq \alpha > 1, \quad j \geq 1, \tag{29}$$

then

$$r_\varepsilon(\mathcal{W}_2^\beta(L)) = r_\varepsilon^{\mathcal{L}}(\mathcal{W}_2^\beta(L)) \left(1 + O\left(\frac{1}{W}\right)\right), \tag{30}$$

where the bandwidth W is defined by (20).

Proof. Let us simplify the lower bound for the minimax risk (24). First of all note that

$$\begin{aligned}
\sum_{j\neq 0}(\theta_j^0)^2 &= \varepsilon^2 \sum_{j\neq 0}\sigma_j^2 \left[1 - \left|\frac{j}{W}\right|^\beta\right]_+ \left|\frac{W}{j}\right|^\beta \\
&= \varepsilon^2 \sum_{j\neq 0}\sigma_j^2 \left[1 - \left|\frac{j}{W}\right|^\beta\right]_+ \left(1 + \left|\frac{j}{W}\right|^\beta - 1\right)^{-1} \\
&\geq \varepsilon^2 \sum_{j\neq 0}\sigma_j^2 \left[1 - \left|\frac{j}{W}\right|^\beta\right]_+ - \varepsilon^2 \sum_{j\neq 0}\sigma_j^2 \left[1 - \left|\frac{j}{W}\right|^\beta\right]_+^2 .
\end{aligned} \tag{31}$$

In order to estimate from above the last term in the right-hand side of the above equation we use (29) and the inequality $1-(1-x)^\beta \leq \max(1,\beta)x$, which is valid for $0 \leq x \leq 1$. Then we have

$$\begin{aligned}
\sum_{j\neq 0}\sigma_j^2 \left[1 - \left|\frac{j}{W}\right|^\beta\right]_+^2 &= 2\sum_{k=0}^{\lfloor W\rfloor}\sigma_{\lfloor W\rfloor - k}^2 \left[1 - \left(\frac{\lfloor W\rfloor - k}{W}\right)^\beta\right]^2 \\
&\leq 2(1+\beta)^2 W^{-2}\sigma_{\lfloor W\rfloor -1}^2 \sum_{k=1}^{\infty}\alpha^{-2(k-1)}k^2 = C_\alpha \sigma_{\lfloor W\rfloor -1}^2 W^{-2}.
\end{aligned} \tag{32}$$

On the other hand it is evident that

$$\begin{aligned}
r_\varepsilon^{\mathcal{L}}(\mathcal{W}_2^\beta(L)) &= \varepsilon^2 \sum_{j=-\infty}^{\infty}\sigma_j^2 \left[1 - \left|\frac{j}{W}\right|^\beta\right]_+ \\
&\geq 2\varepsilon^2\sigma_{\lfloor W\rfloor -1}^2 \left(1 - \left(\frac{\lfloor W\rfloor - 1}{W}\right)^\beta\right) \geq O\left(\varepsilon^2\sigma_{\lfloor W\rfloor -1}^2 W^{-1}\right).
\end{aligned} \tag{33}$$

Therefore from (32) and (31) we see that

$$\sum_{j\neq 0}(\theta_j^0)^2 \geq r_\varepsilon^{\mathcal{L}}(\mathcal{W}_2^\beta(L)) \left(1 + O\left(\frac{1}{W}\right)\right). \tag{34}$$

Thus to complete the proof it remains to consider the last term in the right-hand side of (24). We have

$$\varepsilon^{-2}\sum_{j\neq 0}(\theta_j^0)^4\sigma_j^{-2} = \varepsilon^2 \sum_{j\neq 0}\sigma_j^2 \left[1 - \left|\frac{j}{W}\right|^\beta\right]_+^2 \left|\frac{W}{j}\right|^{2\beta} \tag{35}$$

$$= \varepsilon^2 \sum_{0<|j|\le W/2} \sigma_j^2 \left[1-\left|\frac{j}{W}\right|^{\beta}\right]_+^2 \left|\frac{W}{j}\right|^{2\beta} + \varepsilon^2 \sum_{|j|>W/2} \sigma_j^2 \left[1-\left|\frac{j}{W}\right|^{\beta}\right]_+^2 \left|\frac{W}{j}\right|^{2\beta}$$

$$\le \varepsilon^2 W^{2\beta} \sum_{0<|j|\le W/2} \sigma_j^2 + 2^{2\beta} \sum_{j\ne 0} \sigma_j^2 \left[1-\left|\frac{j}{W}\right|^{\beta}\right]_+^2 .$$

From (29) we arrive at

$$\varepsilon^2 W^{2\beta} \sum_{0<|j|\le W/2} \sigma_j^2 \le 2(\alpha^2-1)^{-1}\varepsilon^2 \sigma^2_{\lfloor W\rfloor -1} \alpha^{\lfloor W\rfloor +2} W^{2\beta}.$$

Hence from (32), (33) and (35) it follows that

$$\varepsilon^{-2} \sum_{j\ne 0} (\theta_j^0)^4 \sigma_j^{-2} \le O\left(W^{-1} r_\varepsilon^{\mathcal{L}}(\mathcal{W}_2^\beta(L))\right).$$

This inequality together with (34) completes the proof of the theorem.

Our next goal is to specify the asymptotic behaviour of $r_\varepsilon^{\mathcal{L}}(\mathcal{W}_2^\beta(L))$ as $\varepsilon \to 0$.

Proposition 3 *Let (29) is fulfilled, then*

$$r_\epsilon^{\mathcal{L}}(\mathcal{W}_2^\beta(L)) = \frac{L}{(2\pi W_\epsilon)^{2\beta}} + O\left(\frac{1}{W_\epsilon^{2\beta+1}}\right), \tag{36}$$

where W_ϵ is a root of equation

$$\varepsilon^2 W_\varepsilon^{2\beta-1} \sigma^2_{\lfloor W_\varepsilon\rfloor} = CL, \tag{37}$$

here $\lfloor x\rfloor$ means integer part of x and C is an arbitrary positive constant not depending on ε.

Proof. A simple algebra easily reveals (see (19), (20)) that

$$r_\epsilon^{\mathcal{L}}(\mathcal{W}_2^\beta(L)) = \frac{L}{(2\pi W)^{2\beta}} + \varepsilon^2 \sum_{j=-\infty}^{\infty} \sigma_j^2 \left[1-\left|\frac{j}{W}\right|^{\beta}\right]_+^2, \tag{38}$$

where W is defined by (20). At the same time by virtue of (32) and (33)

$$\varepsilon^2 \sum_{j=-\infty}^{\infty} \sigma_j^2 \left[1-\left|\frac{j}{W}\right|^{\beta}\right]_+^2 = O\left(W^{-1} r_\epsilon^{\mathcal{L}}(W_2^\beta(L))\right).$$

Hence we get by (38)

$$r_\epsilon^{\mathcal{L}}(\mathcal{W}_2^\beta(L)) = \frac{L}{(2\pi W)^{2\beta}} + O\left(\frac{1}{W^{2\beta+1}}\right). \tag{39}$$

Thus it remains to obtain a lower and an upper bound for W defined by (20). By (20)

$$\begin{aligned}
L &= \varepsilon^2 \sum_{j=-\infty}^{\infty} \sigma_j^2 \left[1 - \left|\frac{j}{W}\right|^\beta\right]_+ \left|\frac{W}{j}\right|^\beta |2\pi j|^{2\beta} \\
&= 2(2\pi)^{2\beta} W^\beta \varepsilon^2 \sum_{j\geq 0} \sigma_j^2 j^\beta \left[1 - \left|\frac{j}{W}\right|^\beta\right]_+ \\
&= 2(2\pi)^{2\beta} W^\beta \varepsilon^2 \sum_{j=0}^{\lfloor W \rfloor} \sigma_{\lfloor W \rfloor - j}^2 (\lfloor W \rfloor - j)^\beta \left(1 - \left|\frac{\lfloor W \rfloor - j}{W}\right|^\beta\right).
\end{aligned}$$

By Taylor expansion and (29) we get from the above equation that for some constants C_1, C_2 do not depending on ε

$$\begin{aligned}
2(2\pi)^{2\beta} W^\beta \varepsilon^2 \sum_{j\geq 0} \sigma_j^2 j^\beta \left[1 - \left|\frac{j}{W}\right|^\beta\right]_+ &\leq C_1 \varepsilon^2 W^{2\beta-1} \sigma_{\lfloor W \rfloor}^2, \\
2(2\pi)^{2\beta} W^\beta \varepsilon^2 \sum_{j\geq 0} \sigma_j^2 j^\beta \left[1 - \left|\frac{j}{W}\right|^\beta\right]_+ &\geq C_2 \varepsilon^2 W^{2\beta-1} \sigma_{\lfloor W \rfloor - 1}^2.
\end{aligned}$$

Therefore W is located between two roots W_1 and W_2 of the equations

$$\varepsilon^2 W_1^{2\beta-1} \sigma_{\lfloor W_1 \rfloor}^2 C_1 = L, \quad \varepsilon^2 W_2^{2\beta-1} \sigma_{\lfloor W_2 \rfloor - 1}^2 C_2 = L, \tag{40}$$

Taking the logarithm in (40) one concludes that $|W_1 - W_2| < C$. This inequality together with (39) completes the proof.

The following theorem easily follows from Proposition 3 and Theorem 1.

Theorem 2 *Let (29) is fulfilled and W_ε is a root of equation (37). Then*

$$r_\varepsilon(\mathcal{W}_2^\beta(L)) = \frac{L}{(2\pi W_\epsilon)^{2\beta}} + O\left(\frac{1}{W_\epsilon^{2\beta+1}}\right).$$

Let us look how this theorem works in the case of elliptic equation. Remind that in this case

$$\sigma_k^2 = 2(|k|+1)\mathrm{e}^{-2(|k|+1)\log\rho}.$$

Taking the logarithm of (37) we get

$$W_\varepsilon = \tilde{W}_\varepsilon + O(1), \tag{41}$$

where

$$\tilde{W}_\varepsilon = -\frac{1}{2\log\rho}\left(\log\frac{L}{\varepsilon^2} - 2\beta\log\log\frac{L}{\varepsilon^2}\right).$$

Hence

$$\begin{aligned}
r_\varepsilon(\mathcal{W}_2^\beta(L)) &= L\left(-\frac{\log\rho}{\pi\log(L/\varepsilon^2)}\right)^{2\beta} \\
&\quad + 4\beta^2\, L\left(-\frac{\log\rho}{\pi\log(L/\varepsilon^2)}\right)^{2\beta} \frac{\log\log(L/\varepsilon^2)}{\log(L/\varepsilon^2)} + O\left(\log^{-2\beta-1}\frac{L}{\varepsilon^2}\right).
\end{aligned}$$

Thus Theorem 2 gives us the expansion of the minimax risk up to the second order term. The linear estimator

$$\tilde{\theta}_j = \left[1 - \left|\frac{j}{\tilde{W}_\varepsilon}\right|^\beta\right]_+ Y_j, \tag{42}$$

with $\tilde{W}_\varepsilon$ given by (41), is the second order minimax estimator.
Remark 1. The estimator (42) is robust with respect to the constant L in the definition of the class $\mathcal{W}^\beta(L)$: if the true value of L is unknown, but it is known that $0 < L_1 \le L \le L_2 < \infty$ then we can use the new bandwidth

$$\tilde{W}_\varepsilon = -\frac{1}{2\log\rho}\left(\log\frac{\tilde{L}}{\varepsilon^2} - 2\beta\log\log\frac{\tilde{L}}{\varepsilon^2}\right),$$

with an arbitrary $\tilde{L} \in [L_1, L_2]$, and the estimator (42) is again a second order minimax estimator.

In the problem of estimation of the initial condition for the parabolic equation σ_j^2 have the following form (c.f. (11))

$$\sigma_j^2 = \exp(8\pi^2 j^2 T).$$

Therefore taking the logarithm of (37), we get

$$W_\varepsilon^2 = \frac{1}{8\pi^2 T}\left(\log\frac{L}{\varepsilon^2} - \frac{(2\beta-1)}{2}\log\log\frac{L}{\varepsilon^2}\right) + O(1). \tag{43}$$

Hence, by Theorem 2

$$r_\varepsilon(\mathcal{W}_2^\beta(L)) = L\left(\frac{2T}{\log(L/\varepsilon^2)}\right)^\beta + O\left(\log^{-\beta-1/2}\frac{L}{\varepsilon^2}\right).$$

Thus the estimator from (42) with the bandwidth W_ε from (43) is a first order minimax estimator. It is not very difficult to check that the projection estimator

$$\hat{\theta}_j = \begin{cases} Z_j, & |j| \le \hat{W}_\varepsilon, \\ 0, & |j| > \hat{W}_\varepsilon, \end{cases}$$

with

$$\hat{W}_\varepsilon^2 = \frac{1}{8\pi^2 T}\log\frac{L}{\varepsilon^2} - \frac{\gamma}{8\pi^2 T}\log\log\frac{L}{\varepsilon^2},$$

where $\gamma > \beta$ is also the first order minimax estimator.
Remark 2. It is easy to see that the estimator (42) with the bandwidth defined by (43) and the projection estimator $\hat{\theta}_j$ are also robust w.r.t. L in the sense of Remark 1.

Unfortunately Theorem 2 gives us only the first term in the expansion of the minimax risk. This situation is quite different from the elliptic case, where this theorem provides us with the second order term. Therefore, we consider in the next section the parabolic case in more detail.

3 Estimation of initial conditions for parabolic equations

In this section we assume that $\sigma_j^2 = \sigma_{-j}^2$ and

$$\log \sigma_{j+1}^2 - \log \sigma_j^2 = C_\sigma j + o(j), \quad j \to \infty, \tag{44}$$

where $C_\sigma > 0$ is some constant. Define the integer W_ε as follows

$$W_\varepsilon = \arg\min_j \left| \sigma_j^2 (2\pi j)^{2\beta} - L\varepsilon^{-2} \right|. \tag{45}$$

Denote for brevity by bold letters two dimensional vectors, that is $\mathbf{x} = (x_1, x_2)^T$, and $\|\mathbf{x}\|^2 = x_1^2 + x_2^2$. The asymptotic behaviour of the minimax risk is given by the following theorem.

Proposition 4 *Let W_ε be given by (45) and condition (44) be fulfilled. Then, as $\varepsilon \to 0$*

$$\begin{aligned} r_\varepsilon(\mathcal{W}_2^\beta(L)) \;\; = \;\; & \frac{L}{(2\pi W_\varepsilon)^{2\beta}} \inf_{\Psi} \sup_{\|\mathbf{y}\| \le 1} \Big\{ \mathbf{E}\|\Psi(\mathbf{Y}) - \mathbf{y}\|^2 \\ & + \left(\frac{W_\varepsilon}{W_\varepsilon + 1} \right)^{2\beta} (1 - \|\mathbf{y}\|^2) \Big\} + \; O\left(\exp(-CW_\varepsilon)\right), \end{aligned} \tag{46}$$

where

$$\mathbf{Y} = \mathbf{y} + s_\varepsilon \xi, \quad \xi \sim \mathcal{N}(0, E), \quad s_\varepsilon = \frac{\varepsilon \sigma_{W_\varepsilon} (2\pi W_\varepsilon)^\beta}{\sqrt{L}},$$

here $C > 0$ is a certain constant and E is the identity matrix.

Proof. Let us obtain first a lower bound. Assume that only $\bar{\theta}_{W_\varepsilon} = (\theta_{W_\varepsilon}, \theta_{-W_\varepsilon})^T$ and $\bar{\theta}_{W_\varepsilon+1} = (\theta_{W_\varepsilon+1}, \theta_{-W_\varepsilon-1})^T$ are unknown. The others θ_j are assumed to be 0. Then the observations have the form

$$\begin{aligned} \mathbf{Z}_{W_\varepsilon} &= \bar{\theta}_{W_\varepsilon} + \varepsilon \sigma_{W_\varepsilon} \xi_{W_\varepsilon}, \\ \mathbf{Z}_{W_\varepsilon+1} &= \bar{\theta}_{W_\varepsilon+1} + \varepsilon \sigma_{W_\varepsilon+1} \xi_{W_\varepsilon+1}. \end{aligned} \tag{47}$$

Based on these observations we have to estimate $\bar{\theta}_{W_\varepsilon}$ and $\bar{\theta}_{W_\varepsilon+1}$ provided that

$$\|\bar{\theta}_{W_\varepsilon}\|^2 (2\pi W_\varepsilon)^{2\beta} + \|\bar{\theta}_{W_\varepsilon+1}\|^2 (2\pi(W_\varepsilon + 1))^{2\beta} \le L. \tag{48}$$

Introduce the new variables

$$\mathbf{y} = \bar{\theta}_{W_\varepsilon} (2\pi W_\varepsilon)^\beta L^{-1/2}, \quad \mathbf{y}_1 = \bar{\theta}_{W_\varepsilon+1} (2\pi(W_\varepsilon + 1))^\beta L^{-1/2}.$$

Then equation (48) can be rewritten as

$$\|\mathbf{y}\|^2 + \|\mathbf{y}_1\|^2 \le 1$$

and the observations (47) are equivalent to the following ones

$$\begin{aligned}\mathbf{Y} &= \mathbf{y} + s_\varepsilon \xi \\ \mathbf{Y}_1 &= \mathbf{y}_1 + s_\varepsilon \sigma_{W_\varepsilon+1}\sigma_{W_\varepsilon}^{-1}(W_\varepsilon+1)^\beta W_\varepsilon^{-\beta}\xi_1,\end{aligned} \tag{49}$$

where ξ and ξ_1 are two-dimensional independent $\mathcal{N}(0,E)$. Thus assuming that $\mathbf{y}_1 = \mathbf{z}\sqrt{1-\|\mathbf{y}\|^2}$, where z_1, z_2 are independent random variables taking values $-1/\sqrt{2}, 1/\sqrt{2}$ with the probabilities 1/2, we have

$$\begin{aligned} r_\varepsilon(\mathcal{W}_2^\beta(L)) &\geq \inf_{\Psi,\Psi_1}\sup_{\|\mathbf{y}\|^2+\|\mathbf{y}_1\|^2\leq 1}\left\{\frac{L}{(2\pi W_\varepsilon)^{2\beta}}\mathbf{E}\left\|\Psi(\mathbf{Y},\mathbf{Y}_1)-\mathbf{y}\right\|^2\right. \\ &+ \left.\frac{L}{(2\pi(W_\varepsilon+1))^{2\beta}}\mathbf{E}\left\|\Psi_1(\mathbf{Y},\mathbf{Y}_1)-\mathbf{y}_1\right\|^2\right\} \\ &\geq \inf_{\Psi,\Psi_1}\sup_{\|y\|\leq 1}\left\{\frac{L}{(2\pi W_\varepsilon)^{2\beta}}\mathbf{E}\mathbf{E}\left\{w\left(\Psi(\mathbf{Y}',\mathbf{Y}_1')-\mathbf{y}\right)|\mathbf{z}\right\}\right. \\ &+ \left.\frac{L}{(2\pi(W_\varepsilon+1))^{2\beta}}\mathbf{E}\mathbf{E}\left\{\left\|\Psi_1(\mathbf{Y}',\mathbf{Y}_1')-\mathbf{z}\sqrt{1-\|\mathbf{y}\|^2}\right\|^2\middle|\mathbf{z}\right\}\right\},\end{aligned} \tag{50}$$

where the new loss function $w(\mathbf{x}) = \min\{\|\mathbf{x}\|^2, 4\}$, and the observations $\mathbf{Y}'$, $\mathbf{Y}_1'$ are given by

$$\begin{aligned}\mathbf{Y}' &= \mathbf{y} + s_\varepsilon \xi \\ \mathbf{Y}_1' &= \mathbf{z}\sqrt{1-\|\mathbf{y}\|^2} + s_\varepsilon \sigma_{W_\varepsilon+1}\sigma_{W_\varepsilon}^{-1}(W_\varepsilon+1)^\beta W_\varepsilon^{-\beta}\xi_1.\end{aligned} \tag{51}$$

To simplify the right-hand side of (50) first of all note that by (28) and (44)

$$\begin{aligned}&\mathbf{E}\mathbf{E}\left\{\left\|\Psi_1(\mathbf{Y}',\mathbf{Y}_1')-\mathbf{z}\sqrt{1-\|\mathbf{y}\|^2}\right\|^2\middle|\mathbf{z}\right\} \\ &\geq (1-\|\mathbf{y}\|^2)\left(1-\frac{4(1-\|\mathbf{y}\|^2)W_\varepsilon^{2\beta}\sigma_{W_\varepsilon}^2}{s_\varepsilon^2\sigma_{W_\varepsilon+1}^2(W_\varepsilon+1)^{2\beta}}\right) \geq 1-\|\mathbf{y}\|^2+O(\exp(-CW_\varepsilon)).\end{aligned}$$

Hence from (50) we have

$$\begin{aligned} r_\varepsilon(W_2^\beta(L)) &\geq \inf_{\Psi}\sup_{\|\mathbf{y}\|\leq 1}\left\{\frac{L}{(2\pi W_\varepsilon)^{2\beta}}\mathbf{E}\mathbf{E}\{w\left(\Psi(\mathbf{Y}',\mathbf{Y}_1')-\mathbf{y}\right)|\mathbf{z}\}\right. \\ &+ \left.\frac{L}{(2\pi(W_\varepsilon+1))^{2\beta}}\left(1-\|\mathbf{y}\|^2\right)\right\} + O(\exp(-CW_\varepsilon)).\end{aligned} \tag{52}$$

To simplify more the right-hand side of the above equation consider the auxiliary observations

$$\mathbf{Y}_1'' = s_\varepsilon \sigma_{W_\varepsilon+1}\sigma_{W_\varepsilon}^{-1}(W_\varepsilon+1)^\beta W_\varepsilon^{-\beta}\xi_1. \tag{53}$$

It is easy to see that L_1-distance between the corresponding densities of the observations $\mathbf{Y}', \mathbf{Y}_1'$ and $\mathbf{Y}', \mathbf{Y}_1''$ defined by (51) and (53), is sufficiently small. Indeed by Cauchy-Schwartz inequality one obtains

$$\begin{aligned}
&\int\int \left| p_{\mathbf{Y}',\mathbf{Y}'_1}(\mathbf{x}_1,\mathbf{x}_2) - p_{\mathbf{Y}',\mathbf{Y}''_1}(\mathbf{x}_1,\mathbf{x}_2)\right| \, \mathrm{d}\mathbf{x}_1 \, \mathrm{d}\mathbf{x}_2 \\
&\quad = \int\int \left| p_{\mathbf{Y}'}(\mathbf{x}_1) p_{\mathbf{Y}'_1}(\mathbf{x}_2) - p_{\mathbf{Y}'}(\mathbf{x}_1) p_{\mathbf{Y}''_1}(\mathbf{x}_2)\right| \, \mathrm{d}\mathbf{x}_1 \, \mathrm{d}\mathbf{x}_2 \\
&\quad = \int \left| p_{\mathbf{Y}'_1}(\mathbf{x}_2) - p_{\mathbf{Y}''_1}(\mathbf{x}_2)\right| \, \mathrm{d}\mathbf{x}_2 \\
&\quad = \int \left| p^{1/2}_{\mathbf{Y}'_1}(\mathbf{x}_2) - p^{1/2}_{\mathbf{Y}''_1}(\mathbf{x}_2)\right| \left(p^{1/2}_{\mathbf{Y}'_1}(\mathbf{x}_2) + p^{1/2}_{\mathbf{Y}''_1}(\mathbf{x}_2)\right) \, \mathrm{d}\mathbf{x}_2 \\
&\quad \le 2 \left(\int \left(p^{1/2}_{\mathbf{Y}'_1}(\mathbf{x}_2) - p^{1/2}_{\mathbf{Y}''_1}(\mathbf{x}_2)\right)^2 \, \mathrm{d}\mathbf{x}_2 \right)^{1/2} \\
&\quad = 2\sqrt{2} \left(1 - \exp\left(- \frac{(1-\|\mathbf{y}\|^2) W_\varepsilon^2 \sigma^2_{W_\varepsilon}}{8 s_\varepsilon^2 (W_\varepsilon+1)^2 \sigma^2_{W_\varepsilon+1}} \right) \right)^{1/2} \le \exp(-CW_\varepsilon).
\end{aligned}$$

Hence for any estimator $\Psi(\cdot)$

$$\mathbf{E}\,\mathbf{E}\{w\left(\Psi(\mathbf{Y}',\mathbf{Y}'_1) - \mathbf{y}\right) | \, \mathbf{z}\} = \mathbf{E}\,\mathbf{E}\{w\left(\Psi(\mathbf{Y}',\mathbf{Y}''_1) - \mathbf{y}\right) | \, \mathbf{z}\} + O\left(\exp(-CW_\varepsilon)\right).$$

Since the *inf* in (52), which is taken over all mesurable functions $\Psi(\cdot)$, can be replaced by the *inf* taken over $\|\Psi\| \le 1$, then using the above equation and (52) we arrive at the following lower bound

$$\begin{aligned}
r_\varepsilon(\mathcal{W}_2^\beta(L)) \quad \ge \quad & \frac{L}{(2\pi W_\varepsilon)^{2\beta}} \inf_{\Psi} \sup_{\|\mathbf{y}\|\le 1} \left\{ \mathbf{E}\,\|\Psi(\mathbf{Y}) - \mathbf{y}\|^2 + \frac{(1-\|\mathbf{y}\|^2) W_\varepsilon^{2\beta}}{(W_\varepsilon+1)^{2\beta}} \right\} \\
& + O\left(\exp(-CW_\varepsilon)\right).
\end{aligned} \tag{54}$$

In order to prove that this lower bound is attainable consider the following estimator

$$\hat{\theta}_k = \begin{cases} Y_k, & |k| < W_\varepsilon, \\ \Psi_k(Y_k, Y_{-k}), & |k| = W_\varepsilon, \\ 0, & |k| > W_\varepsilon, \end{cases}$$

where $\Psi_k(\cdot,\cdot)$ is an arbitrary function $R^2 \to R^1$. Note that by (44) and (45)

$$L\varepsilon^{-2} \ge 0.5 (2\pi W_\varepsilon)^{2\beta} \sigma^2_{W_\varepsilon} = 0.5 (2\pi W_\varepsilon)^{2\beta} \sigma^2_{W_\varepsilon - 1} \mathrm{e}^{(1+o(1)) C_\sigma W_\varepsilon}.$$

Therefore

$$\varepsilon^2 \sum_{j=0}^{W_\varepsilon - 1} \sigma_j^2 \le (1+o(1)) \varepsilon^2 \sigma^2_{W_\varepsilon - 1} = \mathrm{e}^{-(1+o(1)) C_\sigma W_\varepsilon}.$$

Denote for brevity $\Theta' = \{\theta_j : \, \sum_{|j| = W_\varepsilon, W_\varepsilon+1} \theta_j^2 (2\pi j)^{2\beta} \le L\}$. Then we get

$$\begin{aligned}
r_\varepsilon(\mathcal{W}_2^\beta(l)) &\le \varepsilon^2 \sum_{|j| < W_\varepsilon} \sigma_j^2 + \inf_{\Psi} \sup_{\theta} \left\{ \sum_{|j| > W_\varepsilon} \theta_j^2 + \sum_{|j| = W_\varepsilon} \mathbf{E}(\Psi_j(Y_j, Y_{-j}) - \theta_j)^2 \right\} \\
&= \inf_{\Psi} \sup_{\theta \in \Theta'} \left\{ \sum_{|j| = W_\varepsilon} \mathbf{E}(\Psi_j(Y_j, Y_{-j}) - \theta_j)^2 + \sum_{|j| = W_\varepsilon + 1} \theta_j^2 \right\} + O\left(\mathrm{e}^{-CW_\varepsilon}\right) \\
&= \frac{L}{(2\pi W_\varepsilon)^{2\beta}} \inf_{\Psi} \sup_{\|\mathbf{y}\|\le 1} \left\{ \mathbf{E}\|\Psi(\mathbf{Y}) - \mathbf{y}\|^2 + \frac{W_\varepsilon^{2\beta}(1-\|\mathbf{y}\|^2)}{(W_\varepsilon+1)^{2\beta}} \right\} + O\left(\mathrm{e}^{-CW_\varepsilon}\right).
\end{aligned}$$

Together with (54) this inequality completes the proof of the theorem.

At the first glance it seems that Proposition 4, which specifies the minimax risk expansion up to the exponential term, has very restrictive practical meaning since we have to solve the sufficiently complicated variational problem to calculate the vector-valued function $\Psi^0(\cdot,\cdot)$, at which the *inf* in (46) is attained. Fortunately, it can be done very easily and second order minimax estimators can be found explicitly. Moreover we will see that there exists a so-called exponentially efficient estimator, whose risk approximates the minimax risk up to terms of the order $\exp(-CW_\varepsilon)$.

Theorem 3 *Assume that (44) is fulfilled and W_ε is given by (45), then*

$$\begin{aligned} r_\varepsilon(\mathcal{W}_2^\beta(L)) &= \frac{L}{(2\pi W_\varepsilon)^{2\beta}}\left(\left(\frac{W_\varepsilon}{W_\varepsilon+1}\right)^{2\beta} + 2s_\varepsilon^2\left(1-\left(\frac{W_\varepsilon}{W_\varepsilon+1}\right)^{\beta}\right)^2\right) \\ &\quad + \exp(-CW_\varepsilon). \end{aligned} \tag{55}$$

and the following linear estimator

$$\hat{\theta}_k = \begin{cases} Y_k, & |k| < W_\varepsilon, \\ \left(1-\left(\dfrac{W_\varepsilon}{W_\varepsilon+1}\right)^{\beta}\right)Y_k, & |k| = W_\varepsilon, \\ 0, & |k| > W_\varepsilon \end{cases}$$

is exponentially efficient.

Proof. According to Proposition 4 it suffices to solve the following problem. Assume that we are given the noisy data

$$\mathbf{Y} = \mathbf{y} + s_\varepsilon \xi,$$

where ξ is $\mathcal{N}(0,E)$. Based on $\mathbf{Y}$ we have to estimate the unknown vector $\mathbf{y}$. More precisely, our goal is to calculate the minimax risk $\inf_\Psi \sup_{\|\mathbf{y}\|\le 1}\delta_\varepsilon(\Psi,\mathbf{y})$ up to terms of the order $O(\exp(-CW_\varepsilon))$, where

$$\delta_\varepsilon(\Psi,y) = \mathbf{E}\|\Psi(\mathbf{Y})-\mathbf{y}\|^2 + \left(\frac{W_\varepsilon}{W_\varepsilon+1}\right)(1-\|\mathbf{y}\|^2).$$

To get an upper bound consider the linear estimator $\hat{\mathbf{y}} = h\mathbf{Y}$ with

$$h = 1-\left(\frac{W_\varepsilon}{W_\varepsilon+1}\right)^{\beta}.$$

Then elementary algebra reveals that

$$\begin{aligned} \inf_{\Psi}\sup_{\|\mathbf{y}\|\le 1}\delta_\varepsilon(\Psi,\mathbf{y}) &\le \sup_{\|\mathbf{y}\|\le 1}\left\{(1-h)^2\|\mathbf{y}\|^2 + 2h^2 s_\varepsilon^2 + \left(\frac{W_\varepsilon}{W_\varepsilon+1}\right)^{2\beta}(1-\|\mathbf{y}\|^2)\right\} \\ &= \left(\frac{W_\varepsilon}{W_\varepsilon+1}\right)^{2\beta} + 2s_\varepsilon^2\left(1-\left(\frac{W_\varepsilon}{W_\varepsilon+1}\right)^{\beta}\right)^2. \end{aligned} \tag{56}$$

To get a lower bound assume that $\mathbf{y}$ is the Gaussian two-dimensional vector $\mathcal{N}(0, \sigma^2 E)$. Then we have

$$\begin{aligned}
\inf_{\Psi} \sup_{\|\mathbf{y}\| \le 1} \delta_\varepsilon(\Psi, \mathbf{y}) &\ge \inf_{\Psi} \frac{1}{\mathbf{P}\{\|\mathbf{y}\| \le 1\}} \mathbf{E}\delta_\varepsilon(\Psi, \mathbf{y}) \mathbf{1}\{\|\mathbf{y}\| \le 1\} \qquad (57) \\
&= \inf_{\|\Psi\| \le 1} \frac{1}{\mathbf{P}\{\|\mathbf{y}\| \le 1\}} \mathbf{E}\delta_\varepsilon(\Psi, \mathbf{y}) \mathbf{1}\{\|\mathbf{y}\| \le 1\} \\
&\ge \inf_{\Psi} \mathbf{E}\delta_\varepsilon(\Psi, \mathbf{y}) - \frac{5\mathbf{P}\{\|\mathbf{y}\| > 1\}}{\mathbf{P}\{\|\mathbf{y}\| \le 1\}}.
\end{aligned}$$

Since $\mathbf{y}$ is Gaussian the $\inf_\Psi \mathbf{E}\delta_\varepsilon(\Psi, \mathbf{y})$ is attained when

$$\Psi(\mathbf{Y}) = \frac{\sigma^2}{s_\varepsilon^2 + \sigma^2} \mathbf{y}.$$

Let us chose

$$\sigma^2 = s_\varepsilon^2 \left(1 - \left(\frac{W_\varepsilon}{W_\varepsilon + 1}\right)^\beta\right) \left(\frac{W_\varepsilon}{W_\varepsilon + 1}\right)^{-\beta}.$$

Noting that $s_\varepsilon^2 < 2$ we have on the one hand

$$\mathbf{P}\{\|\mathbf{y}\| > 1\} \le C \exp\left(-\frac{1}{2\sigma^2}\right) \le C \exp\left(-\frac{W_\varepsilon}{2\beta s_\varepsilon^2}\right) \le \exp(-CW_\varepsilon).$$

On the other hand noting that under such choice of σ^2

$$h = \frac{\sigma^2}{s_\varepsilon^2 + \sigma^2},$$

and we obtain that (cf. (56))

$$\inf_{\Psi} \mathbf{E}\delta_\varepsilon(\Psi, \mathbf{y}) = \left(\frac{W_\varepsilon}{W_\varepsilon + 1}\right)^{2\beta} + 2s_\varepsilon^2 \left(1 - \left(\frac{W_\varepsilon}{W_\varepsilon + 1}\right)^\beta\right)^2.$$

This inequality together with (56), (57) reveals that

$$\inf_{\Psi} \sup_{\|\mathbf{y}\| \le 1} \delta_\varepsilon(\Psi, \mathbf{y}) = \left(\frac{W_\varepsilon}{W_\varepsilon + 1}\right)^{2\beta} + 2s_\varepsilon^2 \left(1 - \left(\frac{W_\varepsilon}{W_\varepsilon + 1}\right)^\beta\right)^2 + \exp(-CW_\varepsilon).$$

The proof of the theorem follows now from Proposition 4.

References

[1] Hadamard, J. (1932). *Le probléme de Cauchy et les équations aux derivees partielles linéares hyperboliques.* Paris. Herman et Cie.

[2] Chow, P.-L. Khasminskii, R. (1997). Statistical approach to dynamical inverse problems. *Comm. Math. Phys.* To appear.

[3] Chow, P.-L. Ibragimov, I. Khasminskii, R. (1997). Statistical approach to some inverse problems for linear PDE. Submitted to *PTRF*

[4] Ivanov, B. K. Vasin, B. V. Tanana, V. P. (1978). *Theory of linear ill-posed problems and its applications.* Moscow, Nauka.

[5] Lavrentiev, M. M. Romanov, V. G. and Shishtatskii, S. P. (1980). *Ill-posed problems of mathematical physics and analysis.* Moscow, Nauka.

[6] Petrovskii, N. G. (1950). *Lectures on partial differential equations.* Moscow, Gostekhizdat.

[7] Khalfin, L. A. (1978). The statistical approach to the solution of ill-posed problems of geophysics. *Studies in statistical estimation theory.* Zap. Nauchn. Sem. Leningrad. Otd. Inst. Steklov. (LOMI) **79**, 67–81.

[8] Lattes, R. and Lions, J.L. (1967). *Methode de quasi-reversibilite et applications.* Travoux et Recherche Mathematiques. **15**. Dunod, Paris.

[9] Pinsker, M. S. (1980). Optimal filtering of square integrable signals in Gaussian white noise. *Problems of Information Transmission*, 120–133.

[10] Sullivan, F. O'. (1996). A statistical perspective on ill-posed inverse problems. *Statist. Science.* **1** 502–527.

[11] Sudakov, V. N. Khalfin, L. A. (1964). Statistical approach to ill-posed problems in mathematical physics. *DAN USSR.* **157** 1094–1096.

[12] Tihonov, A. N. Arsenin, V. Ja. (1986). *Methods of solutions of ill-posed problems.* Moscow, Nauka.

[13] Walsh, J. B. (1984). An introduction to stochastic PDE. *Lect. Notes in Math.* **1180,** Springer-Varlag, 265–439.