10-1-2007

Image Reconstruction in Multi-Channel Model Under Gaussian Noise

Veera Holdai
Wayne State University

Alexander Korostelev
Wayne State University

IMAGE RECONSTRUCTION IN MULTI-CHANNEL MODEL UNDER GAUSSIAN NOISE

VEERA HOLDAI and ALEXANDER KOROSTELEV

WAYNE STATE UNIVERSITY

Detroit, MI 48202

Department of Mathematics
Research Report

2007 Series
#10

Image reconstruction in multi-channel model under Gaussian noise

Veera Holdai and Alexander Korostelev
Department of Mathematics, Wayne State University
Detroit, Michigan 48202

November 2007

Abstract

The image reconstruction from noisy data is studied. A nonparametric boundary function is estimated from observations in N independent channels in Gaussian white noise. In each channel the image and the background intensities are unknown. They define a non-identifiable nuisance "parameter" that slows down the typical minimax rate of convergence. The large sample asymptotics of the minimax risk is found and an asymptotically optimal estimator for boundary function is suggested.

Key words and phrases: image reconstruction, boundary function estimation, multi-channel model, minimax rates.

AMS 2000 subject classification: Primary 62G08, Secondary 62G20.

1 Introduction

We study a problem that belongs to the image analysis or reconstruction of images from noisy data. Let us start with a statistical model proposed in [10],

$$Y_i = f(X_i) + \xi_i, \qquad i = 1, \dots, n. \tag{1.1}$$

This is a discrete model with a number of observations n, $n \to \infty$. In this model, f is an unknown "intensity" function that depends on a two-dimensional "input" variable $X_i = (X_{i_1}, X_{i_2})$. We call X_i a design point and we assume that it belongs to the unit square $K = [0,1] \times [0,1]$; Y_i is a real-valued response variable determined by the "intensity" function f and the random noise ξ_i. To ease the presentation, suppose that X_i's are independent and uniformly distributed in K while ξ_i's are conditionally independent, given X_i's, with a normal distribution that has zero expectation and a variance $\sigma^2 > 0$, i.e., $\xi_i \sim \mathcal{N}(0, \sigma^2)$.

An "image" is associated with an unknown domain G inside K, and its complement in K, $K \backslash G$, is associated with a background. Assume that $f(x)$ suffers a jump along the boundary of G, i.e., its values are essentially different over the image and the background,

$$f(x) = \begin{cases} f_1(x) & \text{if } x \in G, \\ f_2(x) & \text{if } x \in K \backslash G. \end{cases}$$

where $x = (x_1, x_2)$. Though the model (1.1) resembles a regression model, the objective is not to estimate f. In image analysis, the goal is to estimate the boundary of G, i.e., the curve of discontinuity of f.

Let us discuss in brief a continuous analogue of the model (1.1) :

$$\dot{Y}(x_1, x_2) = f(x_1, x_2) + \varepsilon \dot{W}(x_1, x_2), \quad (x_1, x_2) \in K, \tag{1.2}$$

where $\dot{W}(x_1, x_2)$ is a two-dimensional white noise - a formal derivative of the two-dimensional Wiener sheet $W(x_1, x_2)$, (see [5] or [10]). A small parameter $\varepsilon > 0$, is connected to the discrete model (1.1) by the equation $\varepsilon = \sigma/\sqrt{n}$.

The easiest way to explain the link between (1.1) and (1.2) is to assume that the design points X_i's are not random but rather run over the uniform equidistant grid of points in the unit square K with the step size $1/\sqrt{n}$ in each dimension. There are $[\sqrt{n}x_1] \times [\sqrt{n}x_2]$ observations in the rectangle $R = [0, x_1] \times [0, x_2]$ where $[\sqrt{n}x]$ is the integer part of $\sqrt{n}x$. Sum up and average the discrete observations Y_i over the rectangle R, and obtain the equation,

$$Y^{(n)}(x_1, x_2) = \frac{1}{n} \sum_{X_i \in R} Y_i = \frac{1}{n} \sum_{X_i \in R} f(X_i) + \frac{1}{\sqrt{n}} \sum_{X_i \in R} \frac{\xi_i}{\sqrt{n}}.$$

As $n \to \infty$, the first deterministic Riemann's sum on the right-hand side converges to $\int_0^{x_1} \int_0^{x_2} f(s_1, s_2)\, ds_2\, ds_1$, while the normalized random sum of ξ_i's converges to a two-dimensional random field $\sigma W(x_1, x_2)$ called the Wiener sheet (of intensity σ^2):

$$\sum_{X_i \in R} \frac{\xi_i}{\sqrt{n}} \to \sigma W(x_1, x_2), \qquad 0 \le x_1, x_2 \le 1,$$

Here $W(x_1, x_2)$ is a standard Weiner sheet - a Gaussian random field that has the zero-mean and whose variance equals the area of the rectangle R. The covariance of this random field is given by the formula

$$\mathbf{E}[W(x_1, x_2) W(x_1', x_2')] \;=\; \min[x_1; x_1'] \, \min[x_2; x_2']. \tag{1.3}$$

A natural analogue to the discrete observations $Y^{(n)}(x_1, x_2)$ above is the random field of the continuous observations $Y(x_1, x_2)$ which satisfies the following equation:

$$Y(x_1, x_2) = \int_0^{x_1} \int_0^{x_2} f(s_1, s_2)\, ds_2\, ds_1 + \varepsilon W(x_1, x_2) \tag{1.4}$$

where a small parameter ε is a substitution for $\sigma n^{-1/2}$. In the model (1.4), the asymptotics is studied as $\varepsilon \to 0$. Another traditional notation for $Y(x_1, x_2)$ in (1.4) is in the differentials,

$$dY(x_1, x_2) \;=\; f(x_1, x_2)\, dx_1\, dx_2 \;+\; \varepsilon\, dW(x_1, x_2)$$

or in the formal derivatives,

$$\dot{Y}(x_1, x_2) = f(x_1, x_2) + \varepsilon \dot{W}(x_1, x_2), \tag{1.5}$$

where $\dot{W}(x_1, x_2)$ is a two-dimensional white noise. The differential representation (1.5) is only a convenient notation. The mathematically rigorous interpretation of such models is possible only in the integral sense (1.4).

Note that a consistent estimation in the model (1.4) is possible due to a small parameter ε. Probably, the first work where a continuous white noise image model has been introduced is [6]. In this paper, the likelihood ratio was found and its asymptotics was studied as $\varepsilon \to 0$ in a parametric image model. In nonparametric problems the statistical models are studies with image domains G or their edges not described by finitely many parameters [10]. In nonparametric problems the key question is about estimators that are uniformly good over a broad classes of domains. One possible approach is in the minimax optimality of estimators. The minimax rates of convergence guarantee a certain degree of approximation for any domain within the given class of domains. In the parametric case, the minimax rates of convergence have been studies for a variety of models [10, 11, 9]. Many works in image analysis are practically motivated, e.g., the deconvolution methods [3, 4], the productivity analysis [2, 9], among others. Adaptive estimation in image reconstruction is another interesting direction. In this case, we deal with many nonparametric models, and we wish to find an estimator which is optimal or near optimal for each model without information about the true model. An example is the estimation of image boundaries of unknown degree of smoothness [1].

It is worthy mentioning a closely related area of studies: estimation of support of a density. This density can be either a probability density or an intensity of a Poisson point process [10, 11]. The minimax approach and the rates of convergence turn out to be quite similar in image and density supports estimation. It is also worth to notice that the one-dimensional analogue of an image estimation problem is a change-point problem. For possible estimators and their rates of convergence we refer to [8], [7] and [12].

2 Multi-channel model

Suppose we have a single observation of an unknown image in (1.4) with $\varepsilon = 1$:

$$Y(x_1, x_2) = \int_0^{x_1} \int_0^{x_2} f(s_1, s_2)\, ds_2\, ds_1 \,+\, W(x_1, x_2)\,.$$

In this case, no consistency in estimation of f can be expected from this observation because of the non-decreasing intensity of noise. But what if we have many such observations? We associate each observation with a "channel" and we interpret a set of such observations as a multi-channel image model:

$$Y_j(x_1, x_2) \;=\; \int_0^{x_1} \int_0^{x_2} f_j(s_1, s_2)\, ds_2\, ds_1 \;+\; W_j(x_1, x_2), \qquad j = 1, \dots, N\,, \tag{2.1}$$

where N is a number of channels. The model (2.1) is the principal object of our study. It describes N independent "snap-shots" of the same unknown image G. The unknown intensity functions f_j may differ in different channels; $W_j(x_1, x_2)$ represents a noise in the channel number j , $j = 1, \dots, N$. The random fields W_j are the independent standard 2-D Wiener sheets.

The general model (2.1) is a challenging one, we do not have its complete solution. To simplify the model, suppose that the functions f_j are piecewise-constants;

$f_j = \theta'_j$ if $(x_1, x_2) \in G$ and $f_j = \theta''_j$ if $(x_1, x_2) \in K \backslash G$ where the real-valued constants θ'_j and θ''_j are unknown.

Starting from now on, we assume that the image domain G can be represented by a real-valued function $x_2 = \tau(x_1)$ so that the image domain G can be presented by

$$G = \{(x_1, x_2): \ x_2 < \tau(x_1)\}.$$

In [10], this case is called a *boundary fragment* and τ is a *boundary function.* The boundary function τ is assumed to be one and the same in all the channels. This function is the objective of the statistical estimation from the observations (2.1).

Introduce a vector of all the unknown constant parameters, $\theta = \{(\theta'_j, \theta''_j), j = 1, \ldots, N\}$, $\theta \in \mathbb{R}^{2N}$. Observe that there can be no jump of f_j in some channels if $\theta'_j = \theta''_j$. We will show that for the consistent estimation of the boundary function the difference $|\theta'_j - \theta''_j|$ must be large in some integrated or "averaged" sense.

For the boundary fragments, our model (2.1) can be written explicitly,

$$Y_j(x_1, x_2) = \int_0^{x_1} \int_0^{x_2} [\,\theta'_j \mathbf{1}(s_2 < \tau(s_1)) + \theta'_j \mathbf{1}(s_2 \geq \tau(s_1))\,] ds_2 \, ds_1 + W_j(x_1, x_2)$$

or in differential form,

$$\dot{Y}_j(x_1, x_2) = \theta'_j \mathbf{1}(x_2 < \tau(x_1)) + \theta''_j \mathbf{1}(x_2 \geq \tau(x_1)) + \dot{W}_j(x_1, x_2), \qquad j = 1, \ldots, N. \tag{2.2}$$

The model (2.2) has a "double" nonparametric structure. First, it has the nonparametric part that comes from the unknown boundary function $\tau(x_1)$. Second, there is a growing number $2N$ of unknown constants θ'_j and θ''_j, and we have to take into account this growing dimension of the "nuisance" parameter θ. Note that the components of θ are not identifiable, i.e., they cannot be estimated consistently as $N \to \infty$. As shown below, the rate of convergence in the boundary function estimation should be associated to the "total jump" - the quadratic norm of jumps:

$$\|\Delta\theta\|^2 = \sum_{j=1}^{N} \Delta\theta_j^2 = \sum_{j=1}^{N} (\theta'_j - \theta''_j)^2. \tag{2.3}$$

The rate of convergence depends as well on the *a priori* degree of smoothness of the boundary function. We work with the Hölder smoothness of an integer degree β, $\beta \in \{1, 2, \ldots\}$.

Definition 2.1. *Let β be an integer and $L > 0$. Let $\Sigma(\beta, L)$ denote all the functions $\tau(x_1)$ whose $(\beta - 1)$-th derivative satisfies the Lipschitz condition:*

$$|\tau^{(\beta-1)}(x_1) - \tau^{(\beta-1)}(x_1 + h)| \le L|h|, \qquad x_1, x_1 + h \in [0,1].$$

Functions in $\Sigma(\beta, L)$ can be unbounded and their values can leave the interval $[0,1]$. We restrict their values to even a shorter interval $[t_0, 1-t_0]$, $0 < t_0 < 1/2$, introducing a prior set of functions,

$$\Sigma(\beta, L, t_0) = \Sigma(\beta, L) \cap \{\tau(x_1) : \ 0 \le x_1 \le 1, \ t_0 \le \tau(x_1) \le 1 - t_0\}.$$

Consider the image domain G,

$$G = \{X = (x_1, x_2) \in K : \ 0 \le x_1 \le 1, \ 0 \le x_2 < \tau(x_1)\}.$$

Let $\hat{G} = G(\hat{\tau})$ be an estimator of the domain G obtained from the observations (2.1). The estimator $G(\hat{\tau})$ will be defined via the corresponding estimator $\hat{\tau}(\cdot)$ of the boundary function, $G(\hat{\tau}) = \{(x_1, x_2) \in K : \ x_2 < \hat{\tau}(x_1)\}$. Note that the estimator $\hat{\tau}(x_1)$, $0 \le x_1 \le 1$, is not necessarily a smooth function.

The notation $\mathbb{E}_{\theta,\tau}[\,\cdot\,]$ will be used for the expectation with respect to the distribution $\mathbb{P}_{\theta,\tau}$ of the observations in (2.2) with a boundary function τ and a given set of constants θ, $\theta \in \mathbb{R}^{2N}$.

Our multi-channel image model (2.2) is the extension of its one dimensional analogue proposed in [8]. In the one dimensional case, the image model turns into the change-point problem. Indeed, let x_1 be fixed so that the boundary function τ shrinks to a single point in the interval $[0,1]$. The intensity function f_j equals, respectively, θ'_j or θ''_j before and after τ if we interpret $t = x_2$ as a time scale. So, the two-dimensional observations (2.2) come down to equations,

$$\dot{Y}_j(t) = \theta'_j 1(t < \tau) + \theta''_j 1(t \ge \tau) + \dot{W}_j(t), \quad 0 \le t \le 1, \qquad j = 1, \dots, N, \quad (2.4)$$

As shown in [8], the rate of estimation of the one dimensional parameter τ from the observations (2.4) depends on the performance of the "total jump" $\|\Delta\theta\|^2$. If this quantity grows slower than $O(\sqrt{N})$ a consistent estimation of τ is not possible. If $\|\Delta\theta\|^2$ increases with N faster than $O(N)$ then the parametric rate of convergence $O(\|\Delta\theta\|^{-2})$ is attainable. Thus, the mostly interesting case - at least theoretically - is under the intermediate conditions,

$$\lim_{N\to\infty} \sqrt{N} / \|\Delta\theta\|^2 = 0, \qquad \lim_{N\to\infty} \|\Delta\theta\|^2/N = 0 \quad \text{as} \quad N \to \infty. \qquad (2.5)$$

In the present study, we always assume that the conditions (2.5) hold. Under these conditions, in the one dimensional case, the minimax rate of convergence has been found. It turns out to be $O(N/\|\Delta\theta\|^4) \to 0$ as $N \to \infty$. As we prove in the next section, in the multi-channel image model, the rate of convergence is also associated with the same quantity though in a more complex way which involves the smoothness parameter β of the boundary function. In Section 4, the minimax lower bound is given which claims that the estimator of Section 3 cannot be improved uniformly over the Hölder class of boundary functions. The proofs of the auxiliary lemmas are postponed to the Appendix.

3 Estimation at a point

Motivation: The case of known θ's. Consider the problem of estimation of the boundary function $\tau(x_1)$ at a single point $x_1 = a$, where a is strictly inside the interval $[0,1]$, i.e., $0 < a < 1$. The main result of this section states that uniformly over the boundary functions τ in $\Sigma(\beta, L, t_0)$, the value $\tau(a)$ can be estimated with the rate $O(\varepsilon_N^{2\beta/(\beta+1)})$ where $\varepsilon_N^2 = N/\|\Delta\theta\|^4$. Recall that under the conditions (2.5), $\varepsilon_N \to 0$ as $N \to \infty$.

We want to use a nonparametric version of the maximum likelihood estimator. To explain the underlying motivations, consider the problem with known θ's. If θ's are known in the model (2.2) then we are in position to introduce a single random field $Z(x_1, x_2)$ by

$$\dot{Z}(x_1, x_2) = \frac{1}{\|\Delta\theta\|^2} \sum_{j=1}^{N} (\theta_j' - \theta_j'')\, \mathbf{1}(x_2 < \hat{\tau}(x_1)) \left[\dot{Y}_j(x_1, x_2) - \frac{1}{2}(\theta_j' + \theta_j'') \right]. \quad (3.1)$$

It will be shown that (3.1) is a likelihood function. Actually, it is a functional as it depends on observations $\dot{Y}_j$ and the whole function $\hat{\tau} = \hat{\tau}(\cdot)$. We emphasize this dependence writing $Z = Z(x_1, x_2 \,|\, \hat{\tau})$. The function $\hat{\tau}(\cdot)$ should be looked at as an "input variable" in $Z(x_1, x_2 \,|\, \hat{\tau})$. To define the maximum likelihood estimator, the maximization over $\hat{\tau}$ of this functional must be specified and explained. This is done below.

To understand what $Z(x_1, x_2 \,|\, \hat{\tau})$ in (3.1) has to do with the likelihood, consider the formal log-likelihood of the Gaussian distribution scaled by $\|\Delta\theta\|^2$:

$$-\frac{1}{2\,\|\Delta\theta\|^2}\sum_{j=1}^{N}\big(\,\dot Y_j(x_1\,,x_2)\;-\;[\,\theta'_j\,\mathbf{1}(x_2<\hat\tau(x_1))\;+\;\theta''_j\,\mathbf{1}(x_2\ge\hat\tau(x_1))\,]\,\big)^2 .$$

Leaving only the terms depending on $\hat\tau$, we obtain the log-likelihood functional,

$$\dot Z_0(x_1,x_2\,|\,\hat\tau) \;=\; \frac{1}{\|\Delta\theta\|^2}\sum_{j=1}^{N}\;\big(\,\dot Y_j(x_1\,,x_2)\,[\,\theta'_j\,\mathbf{1}(x_2<\hat\tau(x_1))\;+\;\theta''_j\,\mathbf{1}(x_2\ge\hat\tau(x_1))\,]\,-$$

$$-\,\frac{1}{2}\,[\,\theta'_j\,\mathbf{1}(x_2<\hat\tau(x_1))\;+\;\theta''_j\,\mathbf{1}(x_2\ge\hat\tau(x_1))\,]^2\,\big) \qquad (3.2)$$

Lemma 3.1. *The functional* $Z(x_1,x_2\,|\,\hat\tau)$ *in (3.1) equals the log-likelihood functional* $Z_0(x_1,x_2\,|\,\hat\tau)$ *in (3.2) up to an additive term which does not depend on* $\hat\tau$ *. Besides, the random field* $Z(x_1,x_2\,|\,\hat\tau)$ *admits the following representation:*

$$\dot Z(x_1,x_2) \;=\; \mathbf{1}(x_2<\hat\tau(x_1))\,\left[\frac{1}{2}\,(\mathbf{1}(x_2<\tau(x_1))-\mathbf{1}(x_2\ge\tau(x_1)))\;+\;\frac{1}{\|\Delta\theta\|}\,\dot W(x_1,x_2)\right]$$

with a new standard Wiener sheet $W(x_1\,,x_2)$ *.*

Unknown θ's. In our nonparametric problem with unknown θ's, it is reasonable to substitute θ's in (3.1) by their estimates. Recall that we cannot estimate θ'_j and θ''_j consistently in any channel. Nevertheless, we can use some inconsistent estimators with finite stochastic errors. Since $\tau\;=\;\tau(x_1)\;\in\;[t_0, 1-t_0]$, a part of the corresponding random fields of observations $Y_j(x_1,x_2)$ can be used to obtain the direct estimates of θ's for each j . We will use the parts of these fields located within the strips $\mathcal{T}_1\;=\;[0,1]\times[0,\,t_0/8]$, $\mathcal{T}_2\;=\;[0,1]\times[t_0/8,\,t_0/4]$, and $\mathcal{T}_3\;=\;[0,\,1]\times[1-t_0/4,\,1-t_0/8]$, $\mathcal{T}_4\;=\;[0,\,1]\times[1-t_0/8,\,1]$. By the simple averaging we obtain the estimates,

$$\hat\theta^{(1)}_j \;=\; 8\,t_0^{-1}\int_{\mathcal{T}_1} dY_j(x_1\,,x_2)\;=\;\theta'_j+\xi'_j\,,\quad \hat\theta^{(2)}_j\;=\;8\,t_0^{-1}\int_{\mathcal{T}_2} dY_j(x_1\,,x_2)\;=\;\theta'_j+\eta'_j\,,$$

$$\hat\theta^{(3)}_j \;=\; 8\,t_0^{-1}\int_{\mathcal{T}_3} dY_j(x_1\,,x_2)\;=\;\theta''_j+\xi''_j\,,\quad \hat\theta^{(4)}_j\;=\;8\,t_0^{-1}\int_{\mathcal{T}_4} dY_j(x_1\,,x_2)\;=\;\theta''_j+\eta''_j\,.$$

where $\xi'_j\,,\xi''_j\,,\eta'_j$ and η''_j are independent normal random variables with zero mean and variance $8/t_0$. The cause to take the four strips is to make the estimates independent. Thus, we have got the two independent estimates of θ'_j and θ''_j with random errors whose variance $8/t_0$ is finite. Now we are ready to mimic the case of known θ's and to combine the observations $\dot Y_j(x_1,x_2)$'s into a singe random field:

$$\dot Z(x_1,x_2\,|\,\hat\tau)\;=\;\mathbf{1}(x_2<\hat\tau(x_1))\sum_{j=1}^{N}\left[\,(\hat\theta^{(1)}_j-\hat\theta^{(3)}_j)\left(\dot Y_j(x_1,x_2)-\frac{1}{2}\,(\,\hat\theta^{(2)}_j\;+\;\hat\theta^{(4)}_j\,)\right)\right]. \qquad (3.3)$$

The random field $Z(x_1, x_2 \,|\, \hat{\tau})$ in (3.3) plays the same role of a log-likelihood functional as the one in (3.1). Note that there is no factor associated with $\|\Delta\theta\|^2$ in the definition (3.3) of $Z(x_1, x_2 \,|\, \hat{\tau})$. It is quite understandable since this quadratic norm is unknown. Before we formulate the result about the asymptotic structure of the random field in (3.3), introduce the σ-algebra $\mathcal{F}_0$ generated by the fields Y_j in the union of the strips $\mathcal{T}_i$, $i = 1, 2, 3, 4$.

Lemma 3.2. *Let for a θ, $\theta \in \mathbb{R}^{2N}$, the conditions (2.5) hold. Then the log-likelihood functional $Z(x_1, x_2 \,|\, \hat{\tau})$ in (3.3) admits the following representation:*

$$\dot{Z}(x_1, x_2 \,|\, \hat{\tau}) = \|\Delta\theta\|^2 \, \mathbf{1}(x_2 < \hat{\tau}(x_1)) \cdot \\ \cdot \left[\frac{1}{2} \mathbf{1}(x_2 < \tau(x_1))\,(1+\alpha_2) - \frac{1}{2} \mathbf{1}(x_2 \geq \tau(x_1))\,(1+\alpha_3) + \epsilon_N \dot{W}(x_1, x_2) \right] \tag{3.4}$$

with $\epsilon_N = \|\Delta\theta\|^{-2} \sqrt{\|\Delta\theta\|^2 (1+\alpha_1) + 16\, N/t_0}$ and a new standard Wiener sheet $W(x_1, x_2)$ where the random variables α_i, $i = 1, 2, 3$, are $\mathcal{F}_0$-measurable and $|\alpha_i| \to 0$ in $\mathbb{P}_{\theta,\tau}$-probability uniformly over $\tau \in \Sigma(\beta, L, t_0)$. Moreover, if Θ_0 is a set of θ's for which the convergence in (2.5) is uniform, then the random variables $|\alpha_i| \to 0$ in $\mathbb{P}_{\theta,\tau}$-probability uniformly over Θ_0 as well.

Remark 3.1. If we neglect the vanishing terms in (3.4), we obtain the asymptotic representation for $Z(x_1, x_2 \,|\, \hat{\tau})$ in (3.3),

$$\dot{Z}(x_1, x_2 \,|\, \hat{\tau}) \approx \|\Delta\theta\|^2 \, \mathbf{1}(x_2 < \hat{\tau}(x_1)) \cdot \\ \cdot \left[\frac{1}{2} \left(\mathbf{1}(x_2 < \tau(x_1)) - \mathbf{1}(x_2 \geq \tau(x_1)) \right) + 4/\sqrt{t_0}\, \varepsilon_N \dot{W}(x_1, x_2) \right]$$

with $\varepsilon_N^2 = N / \|\Delta\theta\|^4$.

Comparing the latter asymptotic representation with that in Lemma 3.1, we see the two differences. There is an additional factor $\|\Delta\theta\|^2$, and there is a different intensity of the stochastic term. Recall that we want to use the log-likelihood for maximization over $\hat{\tau}$. Clearly, the constant factor does not spoil this game. In what concerns the intensity of the stochastic term, indeed, we have to make some extra payment for unknown θ's.

The maximum likelihood estimator. The key difference between the one dimensional multi-channel change point problem in [8] and the image model of observations (2.2) is that the one dimensional maximization of the log-likelihood (3.3) does not require any knowledge of the nuisance parameters θ, $\theta \in \mathbb{R}^{2N}$, while in the image model the rate of growth of $\|\Delta\theta\|^2$ plays the essential role in the definition of estimator. For this reason, we start with the case when this rate of growth is fixed. Take a sequence of positive numbers ε_N such that $\varepsilon_N \to 0$ as $N \to \infty$, and introduce a set

$$\Theta(\varepsilon_N) = \left\{ \theta,\ \theta \in \mathbb{R}^{2N} :\ \frac{1}{2}\frac{\sqrt{N}}{\varepsilon_N} \le \|\Delta\theta\|^2 \le 2\,\frac{\sqrt{N}}{\varepsilon_N} \right\}.$$

Note that for $\theta \in \Theta(\varepsilon_N)$, the inequalities hold,

$$\frac{1}{4}\ \varepsilon_N^2 \le \frac{N}{\|\Delta\theta\|^4} \le 4\ \varepsilon_N^2 \tag{3.5}$$

so that the magnitude of $N/\|\Delta\theta\|^4$ equals $O(\varepsilon_N^2)$ uniformly over the set $\Theta(\varepsilon_N)$. Clearly, on the set $\Theta(\varepsilon_N)$, the first condition in (2.5) holds. To ensure the second condition in (2.5) we require that $\sqrt{N}\,\varepsilon_N \to \infty$ as $N \to \infty$.

We consider the boundary functions τ from the Hölder class $\Sigma(\beta, L, t_0)$. For a given sequence ε_N, an estimator of τ will be defined that guarantees a certain rate of convergence uniformly over $\theta \in \Theta(\varepsilon_N)$ and $\tau \in \Sigma(\beta, L, t_0)$. Introduce a sequence $\delta_N = \varepsilon_N^{2/(\beta+1)}$. Now, when we know the log-likelihood function $Z(x_1, x_2 \mid \hat{\tau})$, we can define the maximum likelihood estimator for the boundary fragment $\tau(x_1)$. First, consider the case of the polynomial boundary functions presented in the following form:

$$\tau(x_1) = \gamma_0 + \frac{1}{1!}\gamma_1\,(x_1 - a) + \cdots + \frac{1}{(\beta-1)!}\gamma_{\beta-1}\,(x_1 - a)^{\beta-1}. \tag{3.6}$$

Put $\gamma = (\gamma_0, \gamma_1, \ldots, \gamma_{\beta-1})$, $\gamma \in \mathbb{R}^\beta$, for the vector of the polynomial coefficients. To define the maximum likelihood estimator on the set of polynomials, we have to take the maximum of $Z(x_1, x_2 \mid \hat{\tau})$ over all the polynomial coefficients γ, $\gamma \in \mathbb{R}^\beta$. To avoid the technical troubles of maximization, we look at the log-likelihood functional on a discrete subset,

$$\Gamma_N = \left\{ \gamma :\ \ \gamma = (\,\delta_N^\beta m_0,\ \delta_N^{\beta-1} m_1,\ \ldots,\ \delta_N^2 m_{\beta-2},\ \delta_N m_{\beta-1}) \right\} \tag{3.7}$$

where $m = (\,m_0, m_1, \ldots, m_{\beta-1})$ is a β-tuple of integers. We think of (3.6) as a Taylor's expansion of the unknown boundary function $\tau(x_1)$ at $x_1 = a$. For any $\tau \in \Sigma(\beta, L, t_0)$, the derivatives of τ and, respectively, its Taylor's coefficients are bounded,

$$|\tau(x_1)| \le 1\,,\ |\tau^{(1)}(x_1)| \le L_1\,,\ |\tau^{(2)}(x_1)| \le L_2\,,\ \ldots\,, |\tau^{(\beta-1)}(x_1)| \le L_{\beta-1}\,,\quad 0 \le x_1 \le 1\,,$$

with some constants $L_1, \ldots, L_{\beta-1}$. Thus, we can take into account only those integers $m_0, m_1, \ldots, m_{\beta-1}$ in (3.7) which are also bounded,

$$0 \le m_0 \le \delta_N^{-\beta}, \; |m_1| \le L_1 \delta_N^{-(\beta-1)}, \ldots, |m_{\beta-1}| \le L_{\beta-1}\delta_N^{-1}. \tag{3.8}$$

Observe that the total number of such polynomials does not exceed $L_0\, \delta_N^{-\beta(\beta+1)/2}$ where the constant $L_0 = L_0(\beta)$ does not depend on N. Denote the set of polynomials (3.6) with the coefficients $\gamma \in \Gamma_N$ satisfying (3.8) by $\mathcal{M}_\beta$. We will select our maximum likelihood estimator from this set $\mathcal{M}_\beta$.

To define the maximum likelihood estimator of the value $\tau(a)$ at the fixed point $x_1 = a$, we need the log-likelihood (3.3) only within the strip

$$S_N(a) = [a - \delta_N, a + \delta_N] \times [t_0, 1 - t_0].$$

The maximum likelihood estimator $\tau_N^* = \tau_N^*(\cdot)$ is defined as the "point" of maximum,

$$\tau_N^*(\cdot) = \arg \max_{\hat\tau(\cdot) \in \mathcal{M}_\beta} \int_{S_N(a)} \dot Z(x_1, x_2 \,|\, \hat\tau)\, dx_2\, dx_1. \tag{3.9}$$

Due to the properties of the Wiener sheet, a unique point of maximum exists with probability 1. We take the value of τ_N^* at $x_1 = a$ for the maximum likelihood estimator of the boundary function $\tau(a)$. Introduce another likelihood function by

$$\mathcal{L} = \mathcal{L}(\hat\tau) = \mathcal{L}(\hat\tau \,|\, \tau, \theta) = \|\Delta\theta\|^{-2} \int_{S_N(a)} [\dot Z(x_1, x_2 \,|\, \hat\tau) - \dot Z(x_1, x_2 \,|\, \tau)] dx_2\, dx_1.$$

Note that $\mathcal{L}$ cannot be used for estimation since it involves the unknown terms τ and θ. On the other hand, this modified likelihood function $\mathcal{L}(\hat\tau)$ differs from the integral one in (3.9) by only terms that do not depend on $\hat\tau$. It immediately implies that the point of maximum $\tau_N^*(a)$ in (3.9) coincides with the point of maximum of $\mathcal{L}(\hat\tau)$. The asymptotic performance of the latter is much easier to study. The next lemma shows that at any $\hat\tau$ fixed, the value of the modified likelihood $\mathcal{L}(\hat\tau)$ has the Gaussian distribution with the explicit formulas for expectation and variance. Introduce the L_1-norm of the difference of $\hat\tau$ and the true boundary function τ reduced to the interval $[a - \delta_N, a - \delta_N]$ by

$$d_1(\hat\tau, \tau) = \int_{a-\delta_N}^{a+\delta_N} |\hat\tau(x_1) - \tau(x_1)|\, dx_1.$$

Lemma 3.3. *Let the assumptions of Lemma 3.2 hold. Then there exist random variables $\alpha_4 = \alpha_4(\hat\tau, \tau, \theta)$ and $\alpha_5 = \alpha_5(\hat\tau, \tau, \theta)$, measurable with respect to the*

σ-algebra $\mathcal{F}_0$, and vanishing in $\mathbb{P}_{\theta,\tau}$-probability uniformly over $\hat{\tau} \in \mathcal{M}_\beta$, $\theta \in \Theta(\varepsilon_N)$, and $\tau \in \Sigma(\beta, L, t_0)$, i.e., $|\alpha_4| \to 0$ and $|\alpha_5| \to 0$ as $N \to \infty$, *such that, conditionally on $\mathcal{F}_0$, the random variable $\mathcal{L} = \mathcal{L}(\hat{\tau}\,|\,\tau, \theta)$ has the Gaussian distribution,*

$$\mathcal{L} \sim \mathcal{N}\big(-\frac{1}{2}(1+\alpha_4)\, d_1(\hat{\tau}, \tau),\ \frac{16}{t_0}(1+\alpha_5)\, \varepsilon_N^2\, d_1(\hat{\tau}, \tau)\big)$$

where $\varepsilon_N^2 = N/\|\Delta\theta\|^4$.

For the true boundary function $\tau(x_1)$ define its approximate Taylor's polynomial $\tau^{(0)}(x_1)$ at $x_1 = a$ by the formula

$$\tau^{(0)}(x_1) = \delta_N^{\beta}\, m^{(0,0)} + \delta_N^{(\beta-1)}\, \frac{m^{(0,1)}}{1!}(x_1 - a) + \cdots + \delta_N\, \frac{m^{(0,\beta-1)}}{(\beta-1)!}(x_1 - a)^{\beta-1} \quad (3.10)$$

where the integers $m^{(0,i)}$ are given by

$$m^{(0,i)} = \left[\delta_N^{i-\beta}\, \tau^{(i)}(a)\right], \qquad i = 0, 1, \ldots, \beta - 1.$$

Define the vector of these integers,

$$m^{(0)} = (m^{(0,0)}, m^{(0,1)}, \ldots, m^{(0,\beta-1)}), \qquad m^{(0)} \in \mathbb{Z}^{\beta}.$$

The Taylor approximation (3.10) is a convenient tool to describe the distance between the actual function $\tau(x_1)$ and its estimator $\hat{\tau}_N(x_1)$. First, we state a trivial result about the distance between a boundary function and its Taylor's approximation.

Lemma 3.4. *Let $\tau \in \Sigma(\beta, L, t_0)$ and let $\tau^{(0)}(x_1)$ be its approximate Taylor polynomial defined by (3.10). Then there exists a constant $C_T > 0$ which depends only on the class $\Sigma(\beta, L, t_0)$ such that*

$$d_1(\tau, \tau^{(0)}) \le C_T\, \delta_N^{\beta+1} = C_T\, \varepsilon_N^2.$$

The following theorem describes the rate of convergence of the maximum likelihood estimator $\tau_N^*(a)$.

Theorem 3.1. *Uniformly in $\tau \in \Sigma(\beta, L, t_0)$ and $\theta \in \Theta(\varepsilon_N)$, the normalized deviations $\varepsilon_N^{-2\beta/(\beta+1)}\,|\tau_N^*(a) - \tau(a)|$ of the maximum likelihood estimator (3.9) are bounded in $\mathbb{P}_{\theta,\tau}$-probability, that is*

$$\lim_{x\to\infty} \limsup_{N\to\infty} \sup_{\tau\in\Sigma(\beta,L,t_0)} \sup_{\theta\in\Theta(\varepsilon_N)} \mathbb{P}_{\theta,\tau}\Big(\big(\|\Delta\theta\|^4/N\big)^{\beta/(\beta+1)}\, |\tau_N^*(a) - \tau(a)| \ge x \Big) = 0. \quad (3.11)$$

Proof. Choose a large positive number z and take an integer k, $k = 0, 1, \ldots$. For every k, define a "spherical layer" by

$$L_k = \left\{ \hat{\tau}, \hat{\tau} \in \mathcal{M}_\beta : k\, z\, \delta_N^{\beta+1} \leq d_1(\hat{\tau}, \tau^{(0)}) \leq (k+1)\, z\, \delta_N^{\beta+1} \right\}, \tag{3.12}$$

where $\tau^{(0)}$ is the approximate Taylor polynomial of a true boundary function τ. The next auxiliary result estimates the number of the integer points in each layer.

Lemma 3.5. *For any true boundary function τ, the number ν_k, $k \geq 1$, of the elements in L_k does not exceed $A\,(kz)^\beta$ with a positive constant A independent of z or k.*

Return to the proof of the theorem. The idea of the proof is standard for the maximum likelihood estimators. In accordance with Lemma 3.3, the random variable $\mathcal{L}$ has a negative expected value proportional to $d_1(\hat{\tau}, \tau)$ and a variance with a small factor ε_N^2. Consider the random variable $\mathcal{L}(\tau^{(0)} \mid \tau, \theta)$ with $\hat{\tau} = \tau^{(0)}$. By Lemma 3.4, the distance $d_1(\tau^{(0)}, \tau) = O(\varepsilon_N^2)$ is small. We want to show that with a high probability, the random variable $\mathcal{L}(\tau^{(0)} \mid \tau, \theta)$ is bigger than $-z\,\varepsilon_N^2$ for some large z, $z > 0$. On the other hand, if $\hat{\tau}$ belongs to a "spherical layer" L_k with a large k, $k > k_0$, where k_0 is fixed, then the distance $d_1(\hat{\tau}, \tau)$ is large, and the probability of the random event $\mathcal{L}(\hat{\tau} \mid \tau, \theta) \geq -z\,\varepsilon_N^2$ is vanishing as $z \to \infty$. Thus, the "point" of maximum, τ_N^*, with a high probability must belong to one of the "spherical layers" L_k with $k \leq k_0$ and z large enough. It implies that with a high probability the distance $d_1(\tau_N^*, \tau)$ has the magnitude $O(\varepsilon_N^2) = O(\delta_N^{\beta+1})$. Finally, as the following lemma shows, if the distance $d_1(\tau_N^*, \tau)$ has the magnitude $O(\delta_N^{\beta+1})$, then the absolute deviation $|\,\tau_N^*(a) - \tau(a)\,|$ at $x_1 = a$ has the magnitude $O(\delta_N^\beta)$.

Lemma 3.6. *Let $\hat{\tau} \in \mathcal{M}_\beta$ and $\tau \in \Sigma(\beta, L, t_0)$. Let for a given constant z_0 the inequality holds, $d_1(\hat{\tau}, \tau) \leq z_0\, \delta_N^{\beta+1}$. Then there exists a constant $C_0 = C_0(z_0, \beta)$ independent of N and such that $|\,\hat{\tau}(a) - \tau(a)\,| \leq C_0\, \delta_N^\beta$.*

Now we are ready to proceed to the program announced above. Introduce the random event $\mathcal{A}_0 = \{\, |\,\alpha_4\,| \leq 1/2;\ |\,\alpha_5\,| \leq 1/2 \,\}$. From Lemma 3.3, we find that $\mathbb{P}_{\theta,\tau}(\mathcal{A}_0) \to 1$ as $N \to \infty$. Conditionally on $\mathcal{F}_0$, the random variable $\mathcal{L}(\tau^{(0)} \mid \tau, \theta)$

is Gaussian, so that

$$\liminf_{N\to\infty} \mathbb{P}_{\theta,\tau}(\mathcal{L}(\tau^{(0)} \,|\, \tau, \theta) \geq \; - z\, \varepsilon_N^2) =$$
$$= \liminf_{N\to\infty} \mathbb{E}_{\theta,\tau}[\mathcal{A}_0 \, \mathbb{P}_{\theta,\tau} \left(\mathcal{L}(\tau^{(0)} \,|\, \tau, \theta) \geq \; - z\, \varepsilon_N^2 \,|\, \mathcal{F}_0 \right)]$$

For any $\hat{\tau} \in \mathcal{M}_\beta$, put

$$\mu_N(\hat{\tau}) = \frac{1}{2}(1+\alpha_4)\, d_1(\hat{\tau}, \tau) \quad \text{and} \quad \sigma_N(\hat{\tau}) = \Big[\frac{16}{t_0}(1+\alpha_5)\, \varepsilon_N^2\, d_1(\hat{\tau}, \tau) \Big]^{1/2}.$$

Take $\hat{\tau} = \tau^{(0)}$, and compute the conditional probability,

$$\mathbb{P}_{\theta,\tau} \left(\mathcal{L}(\tau^{(0)} \,|\, \tau, \theta) \geq \; - z\, \varepsilon_N^2 \,|\, \mathcal{F}_0 \right) =$$
$$= \mathbb{P}_{\theta,\tau} \Big(\frac{\mathcal{L}(\tau^{(0)} \,|\, \tau, \theta) + \mu_N(\tau^{(0)})}{\sigma_N(\tau^{(0)})} \geq \frac{-z\, \varepsilon_N^2 + \mu_N(\tau^{(0)})}{\sigma_N(\tau^{(0)})} \,|\, \mathcal{F}_0 \Big) = \Phi\Big(\frac{z\, \varepsilon_N^2 - \mu_N(\tau^{(0)})}{\sigma_N(\tau^{(0)})} \Big)$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. In accordance with Lemma 3.4, on the random event $\mathcal{A}_0$, the inequality holds

$$\frac{z\, \varepsilon_N^2 - \mu_N(\tau^{(0)})}{\sigma_N(\tau^{(0)})} \geq \frac{z\, \varepsilon_N^2 - (3/4)\, C_T\, \varepsilon_N^2}{\left[(24/t_0)\, C_T\, \varepsilon_N^4 \right]^{1/2}} \geq c_1\, (z - C_T)$$

with $c_1 = \sqrt{t_0/(24\, C_T)}$. Thus,

$$\liminf_{N\to\infty} \mathbb{E}_{\theta,\tau}[\mathcal{A}_0 \, \mathbb{P}_{\theta,\tau} \left(\mathcal{L}(\tau^{(0)} \,|\, \tau, \theta) \geq \; - z\, \varepsilon_N^2 \,|\, \mathcal{F}_0 \right)] \geq \Phi\left(c_1\, (z - C_T) \right)$$

which implied that

$$\lim_{z\to\infty} \liminf_{N\to\infty} \mathbb{P}_{\theta,\tau}(\mathcal{L}(\tau^{(0)} \,|\, \tau, \theta) \geq \; - z\, \varepsilon_N^2) = 1. \tag{3.13}$$

We want to show that uniformly over θ and τ,

$$\lim_{z\to\infty} \limsup_{N\to\infty} \mathbb{P}_{\theta,\tau} \left(\cup_{k \geq k_0} \cup_{\hat{\tau} \in L_k} \{ \mathcal{L}(\hat{\tau} \,|\, \tau, \theta) \geq \; - z\, \varepsilon_N^2 \} \right) = 0. \tag{3.14}$$

Conditionally, given $\mathcal{F}_0$, if the random event $\mathcal{A}_0$ occurs then

$$\mathbb{P}_{\theta,\tau} \left(\mathcal{L}(\hat{\tau} \,|\, \tau, \theta) \geq \; - z\, \varepsilon_N^2 \,|\, \mathcal{F}_0 \right) = 1 - \Phi\Big(\frac{\mu_N(\hat{\tau}) - z\, \varepsilon_N^2}{\sigma_N(\hat{\tau})} \Big).$$

For any $\hat{\tau} \in L_k$, we have that

$$\mu_N(\hat{\tau}) = \frac{1}{2}(1+\alpha_4)\, d_1(\hat{\tau}, \tau) \geq \frac{1}{4} \left[d_1(\hat{\tau}, \tau^{(0)}) - d_1(\tau^{(0)}, \tau) \right] \geq \frac{1}{4}(k\, z - C_T)\, \varepsilon_N^2$$

and

$$\sigma_N^2(\hat{\tau}) = \frac{16}{t_0}(1+\alpha_5)\, \varepsilon_N^2\, d_1(\hat{\tau}, \tau) \leq \frac{24}{t_0} [(k+1)\, z + C_T]\, \varepsilon_N^4.$$

If $z \ge 1$ and $k \ge k_0 \ge 8 + 2\,C_T$, then

$$\frac{\mu_N(\hat\tau) \,-\, z\,\varepsilon_N^2}{\sigma_N(\hat\tau)} \ge \frac{1}{8}\sqrt{\frac{t_0}{6}}\;\frac{k\,z \,-\, 4\,z \,-\, C_T}{\sqrt{k\,z + z + C_T}} \ge \frac{1}{8}\sqrt{\frac{t_0\,k\,z}{6}}\;\frac{1 - (4 + C_T)/k}{\sqrt{1 + (1 + C_T)/k}} \ge a\sqrt{k\,z}$$

with $a \,=\, \sqrt{2\,t_0}/24$. Hence

$$\mathbb{P}_{\theta,\tau}\left(\mathcal{L}(\hat\tau\,|\,\tau,\theta) \ge \,-z\,\varepsilon_N^2\,|\,\mathcal{F}_0\right) \,\le\, 1 \,-\, \Phi\big(a\sqrt{k\,z}\big)\,.$$

If z is so large that $a\sqrt{k\,z} \,\ge\, 1$, then the elementary inequality $1 \,-\, \Phi(x) \,\le\, \exp(-x^2/2)$, $x \ge 1$, implies

$$\mathbb{P}_{\theta,\tau}\left(\,\cup_{k\ge k_0}\,\cup_{\hat\tau\in L_k}\,\{\mathcal{L}(\hat\tau\,|\,\tau,\theta) \,\ge\, -z\,\varepsilon_N^2\,\}\,\right) \,\le\, \sum_{k=k_0}^{\infty} \nu_k\,\exp(-\frac{a^2\,k\,z}{2}) \le$$

$$\le \sum_{k=k_0}^{\infty} A\,(k\,z)^\beta\,\exp\left(-\,\frac{a^2 k\,z}{2}\right) \,=\, A\,z^\beta\,\exp(-\frac{a^2\,k_0\,z}{2})\sum_{k=k_0}^{\infty} k^\beta\,\exp(-\frac{a^2\,(k-k_0)\,z}{2})\,,$$

where Lemma 3.5 has been applied. The latter infinite sum is finite,

$$\sum_{k=k_0}^{\infty} k^\beta\,\exp(-\frac{a^2\,(k-k_0)\,z}{2}) \le \sum_{k=k_0}^{\infty} k^\beta\,\exp(-\frac{a^2\,(k-k_0)}{2}) \,=\, C_1 \,<\infty\,.$$

Finally, combining these estimates, we find that the limit in (3.14) is zero,

$$\lim_{z\to\infty}\,\limsup_{N\to\infty}\,\mathbb{P}_{\theta,\tau}\left(\,\cup_{k\ge k_0}\,\cup_{\hat\tau\in L_k}\,\{\mathcal{L}(\hat\tau\,|\,\tau,\theta) \,\ge\, -z\,\varepsilon_N^2\,\}\,\right) \,\le$$

$$\le \lim_{z\to\infty}\,C_1\,A\,z^\beta\,\exp(-\frac{a^2\,k_0\,z}{2}) \,=\, 0\,.$$

The interpretation of the inequalities (3.13) and (3.14) is immediate. For an arbitrarily small p, $p \,>\, 0$, there exists a positive number z_0 such that with the probability at least $1-p$, the maximum likelihood estimator τ_N^* belongs to one of the layers L_k with $k \le k_0$ and $z = z_*$. In its turn, this fact and the definition of the layers L_k guarantee that $d_1(\tau_N^*\,,\,\tau^{(0)}) \le (k_0+1)\,z_*\,\delta_N^{\beta+1}$. Applying Lemma 3.6 with $z_0 \,=\, (k_0+1)\,z_*$, we obtain the inequality

$$\limsup_{x\to\infty}\,\limsup_{N\to\infty}\,\mathbb{P}_{\theta,\tau}\left(\,\varepsilon_N^{-2\,\beta/(\beta+1)}\,|\,\tau_N^*(a) \,-\, \tau(a)\,| \ge x\right) \,\le$$

$$\le \limsup_{x\to\infty}\,\mathbb{P}_{\theta,\tau}\left(\,|\,\tau_N^*(a) \,-\, \tau(a)\,| \ge C_*\,\varepsilon_N^{2\,\beta/(\beta+1)}\right) \le p\,.$$

Since p is arbitrarily small, the limit as $x \to \infty$ on the left-hand side of (3.11) exists and equals zero. This proves the theorem. □

Adaptation to unknown θ. As mentioned above, Theorem 3.1 guarantees the rate of convergence $\varepsilon_N^{2\,\beta/(\beta+1)}$ uniformly over all the boundary functions $\tau \in \Sigma(\beta\,,L\,,t_0)$

but only locally in θ. Under our assumption, θ must belong to $\Theta(\varepsilon_N)$ defined by a chosen sequence ε_N. It is understandable since the bandwidth δ_N is defined in terms of this sequence. Now, the question arises: Is it possible to substitute an unknown $\varepsilon_N^2 = N/\|\Delta\theta\|^4$ by an estimate obtainable from observation (2.2) so that an analogue of Theorem 3.1 would stay valid uniformly over all θ's ? We will show that under some restrictions, the answer to this question is positive.

Let Θ_0 be a set of θ's, $\theta \in \mathbb{R}^{2N}$, for which the convergence in (2.5) is uniform. An example of such a set Θ_0 can be presented by

$$\Theta_0 = \{ \theta : \psi_N^{-1} \sqrt{N} \leq \|\Delta\theta\|^2 \leq \psi_N N \}$$

where a sequence of positive numbers ψ_N is given in advance, $\psi_N \to 0$ as $N \to \infty$. Note that in this example, the sequence ψ_N can decrease in whatever slow rate. There is a strip in the unit square K not yet used in our considerations, $\mathcal{T}_0 = [0,1] \times [t_0/2, t_0]$. Take $\hat{\tau} = 1$ in (3.3), and define a statistic totally computable from the observations (2.2) :

$$Q_N^* = \frac{4}{t_0} \int_{\mathcal{T}_0} \dot{Z}(x_1, x_2 \,|\, \hat{\tau} = 1)\, dx_2\, dx_1.$$

The following lemma is the immediate consequence of Lemma 3.3.

Lemma 3.7. *Conditionally on $\mathcal{F}_0$, the random variable Q_N^* is Gaussian,*

$$Q_N^* \sim \mathcal{N}\left(\|\Delta\theta\|^2 (1 + \alpha_2),\ \frac{8}{t_0} \epsilon_N^2 \|\Delta\theta\|^4 \right)$$

and $Q_N^* = \|\Delta\theta\|^2 (1 + \alpha_6)$ where $\alpha_6 \to 0$ in $\mathbb{P}_{\theta,\tau}$-probability uniformly over $\tau \in \Sigma(\beta, L, t_0)$, and $\theta \in \Theta_0$.

The statistic Q_N^* can serve as a empirical substitution for $\|\Delta\theta\|^2$. As a Gaussian random variable, it may take negative values and values whatever close to zero. That is why, we replace Q_N^* by a truncation to define an empirical analogue of the bandwidth $\delta_N = \varepsilon_N^{2/(\beta+1)}$ by

$$\delta_N^* = \min \left[\left(N/(Q_N^*)^2\right)^{1/(\beta+1)} ;\ a;\ 1 - a \right]$$

and put

$$S_N^*(a) = [a - \delta_N^*, a + \delta_N^*] \times [t_0, 1 - t_0]. \tag{3.15}$$

The truncation in the definition of δ_N^* guarantees that the strip $S_N^*(a)$ lies entirely within the unit square with probability 1. The strip $S_N^*(a)$ serves as a substitution for $S_N(a)$ which is not computable from the data in the multi-channel model. It is worthy to mention that δ_N^* and the strip $S_N^*(a)$ are $\mathcal{F}_0$-measurable while the white noise in (2.2), $\{\, \dot{W}_j(x_1\,,x_2)\,,\ \ (x_1\,,x_2)\, \in\, S_N^*(a)\,\}$ is independent of $\mathcal{F}_0$.

Theorem 3.2. *Substitute δ_N in the definition of the maximum likelihood estimator τ_N^* by δ_N^*. Then the statement of Theorem 3.1 stays true uniformly over $\tau \in \Sigma(\beta\,,L\,,t_o)$ and $\theta \in \Theta_0$.*

4 Lower bound

Recall, our model is (2.2):

$$\dot{Y}_j(x_1,x_2) = \theta_j' \mathbf{1}(x_2 < \tau(x_1)) + \theta_j'' \mathbf{1}(x_2 \geq \tau(x_1)) + \dot{W}_j(x_1,x_2)$$

where $\theta = \{(\theta_j', \theta_j''), j = 1, \ldots, N\} \in \mathbb{R}^{2N}$ is unknown.

As in the upper bound, take a sequence of positive numbers ε_N such that

$$\varepsilon_N \to 0 \qquad \text{and} \qquad \varepsilon_N \sqrt{N} \to \infty \quad \text{as} \quad N \to \infty\,. \tag{4.1}$$

Introduce a prior set

$$\Theta(\varepsilon_N) = \left\{ \theta \,:\, \theta \,\in\, \mathbb{R}^{2N} \,:\, \frac{1}{2}\, \frac{\sqrt{N}}{\varepsilon_N} \leq \|\Delta\theta\|^2 \leq 2\, \frac{\sqrt{N}}{\varepsilon_N} \right\} \tag{4.2}$$

and choose $\delta_N \;=\; \varepsilon_N^{2/(\beta+1)}$.

Theorem 4.1. *There exists a positive constant C_L such that for any sequence ε_N satisfying (4.1) and the sequence δ_N defined above the lower bound holds:*

$$\liminf_{N\to\infty} \inf_{\hat{\tau}_N} \sup_{\tau\,,\theta\in\Theta(\varepsilon_N)} \delta_N^{-\beta}\, \mathbb{E}_{\theta\,,\tau}\, |\, \hat{\tau}_N(a) - \tau(a)\,| \,\geq\, C_L, \tag{4.3}$$

where the prior set $\Theta(\varepsilon_N)$ is defined by (4.2).

Proof. Consider a subset $\Theta^{(0)}(\varepsilon_N)$ of $\Theta(\varepsilon_N)$ defined by

$$\Theta^{(0)}(\varepsilon_N) = \left\{ \theta \,=\, (\,\theta_j' \,=\, \Delta\theta_j,\; \theta_j'' = 0\,)_{\,j=1}^{\,N} \,:\, \frac{\sqrt{N}}{2\,\varepsilon_N} \,\leq\, \|\,\Delta\theta\|^2 \,\leq\, \frac{2\sqrt{N}}{\varepsilon_N} \right\}.$$

By $\mathbb{E}_{\theta\,,\tau}$ and $\mathbb{P}_{\theta\,,\tau}$ we understand the expectation and distribution of observations in (2.2) for a given $\tau(x_1)$ and $\theta = \{(\Delta\theta_j, 0), j = 1, \ldots, N\}$.

Put $\sigma^2 \,=\, \sigma_N^2 \,=\, 1/(\varepsilon_N\sqrt{N})$ and introduce a sequence $\Delta\theta_j$ of the independent normal random variables with zero-mean and variance σ^2.

Denote by $\mathbb{E}^{(\Delta\theta)}$ the expectation with respect to the distribution $\mathbb{P}^{(\Delta\theta)}$ of these normal random variables. Note that for N large the following lower bound is true

$$\mathbb{P}^{(\Delta\theta)}\left(\theta \in \Theta^{(0)}(\varepsilon_N)\right) = \mathbb{P}^{(\Delta\theta)}\left(\frac{1}{2}\frac{\sqrt{N}}{\varepsilon_N} \le \sum_{j=1}^{N} \Delta\theta_j^2 \le 2\frac{\sqrt{N}}{\varepsilon_N}\right) =$$
$$= \mathbb{P}^{(\Delta\theta)}\left(\frac{1}{2} \le \frac{\sum_{j=1}^{N}\Delta\theta_j^2}{N\sigma^2} \le 2\right) \ge 1 - e^{-N}, \qquad (4.4)$$

where $(\sum_{j=1}^{N}\Delta\theta_j^2)/(N\sigma^2)$ is a standard chi-square random variable with N degrees of freedom. Next, we will need two hypotheses on function $\tau(x_1)$. Let $\tau_0(x_1) = C$ for $x_1 \in [0,1]$ with a constant $C \in (t_0, 1-t_0)$ and let the other hypothesis differ from $\tau_0(x_1)$ by a "bump" of height δ_N^β centered at a point $a \in (0,1)$, i.e. let $\tau_1(x_1) = C + \delta_N^\beta\varphi_0(\frac{x_1-a}{\delta_N})$, where $\varphi_0 > 0$ is some test function, such that $\varphi_0 \in \Sigma(\beta, L, t_0)$.

Notice that both, $\tau_0(x_1)$ and $\tau_1(x_1)$ belong to our class of boundary functions $\Sigma(\beta, L, t_0)$. Indeed, $\tau_0(x_1) \in \Sigma(\beta, L, t_0)$ trivially. By definition of $\Sigma(\beta, L, t_0)$ if $\varphi_0 \in \Sigma(\beta, L, t_0)$, then

$$\left|\varphi_0^{\beta-1}(x_1) - \varphi_0^{\beta-1}(x_1')\right| \le L|x_1 - x_1'|.$$

For $\varphi(x_1) = \delta_N^\beta\varphi_0(\frac{x_1-a}{\delta_N})$ this gives the following:

$$\left|\varphi^{(\beta-1)}(x_1) - \varphi^{(\beta-1)}(x_1')\right| =$$
$$= \left|\delta_N^\beta\,\delta_N^{-(\beta-1)}\varphi_0^{(\beta-1)}\left(\frac{x_1-a}{\delta_N}\right) - \delta_N^\beta\,\delta_N^{-(\beta-1)}\varphi_0^{(\beta-1)}\left(\frac{x_1'-a}{\delta_N}\right)\right| \le$$
$$\le \delta_N \cdot L\left|\frac{x_1-a}{\delta_N} - \frac{x_1'-a}{\delta_N}\right| = L\cdot|x_1 - x_1'|.$$

Hence, $\tau_1(x_1) = C + \delta_N^\beta\varphi_0(\frac{x_1-a}{\delta_N}) \in \Sigma(\beta, L, t_0)$.

For the two hypotheses $\tau_i(x_1)$, $i = 0, 1$ denote the sets

$$G_i = \{(x_1, x_2) : \le x_1 \le 1, 0 \le x_2 \le \tau_i(x_1)\},$$

and by $S_i = S(G_i)$, $i = 0, 1$ denote the corresponding areas.

We impose the following condition on our hypotheses: we require that the difference between the areas S_1 and S_0 be of order ε_N^2, i.e. $S_1 - S_0 = r_0\,\varepsilon_N^2$. This is the least difference when we can distinguish our two hypotheses. The positive constant r_0 will guarantee the right constant L in the class $\Sigma(\beta, L, t_0)$.

Our goal is to prove the lower bound for an estimator of the true function $\tau(x_1)$ at a fixed point $a \in (0,1)$. For any estimator $\hat{\tau}_N(a)$ the maximum of the expected losses can be estimated from below by the Bayes risk:

$$\sup_{\tau\in\Sigma(\beta,L,t_0),\,\theta\in\Theta^{(0)}(\varepsilon_N)} \delta_N^{-\beta}\,\mathbb{E}_{\theta,\tau}\,|\,\hat\tau_N(a)\,-\,\tau(a)\,|\,\geq$$

$$\geq \sup_{\tau\in\{\tau_0,\tau_1\},\,\theta\in\Theta^{(0)}(\varepsilon_N)} \delta_N^{-\beta}\,\mathbb{E}_{\theta,\tau}\,|\,\hat\tau_N(a)\,-\,\tau(a)\,|\,\geq$$

$$\geq \delta_N^{-\beta}\,\mathbb{E}^{(\Delta\theta)}\left[\mathbf{1}\left\{\theta\in\Theta^{(0)}(\varepsilon_N)\right\}\left(\frac{1}{2}\,\mathbb{E}_{\theta,\tau_0}\,|\,\hat\tau_N(a)\,-\,\tau_0(a)\,|\,+\,\frac{1}{2}\,\mathbb{E}_{\theta,\tau_1}\,|\,\hat\tau_N(a)\,-\,\tau_1(a)\,|\right)\right]\geq$$

$$\geq \frac{1}{4}\,\mathbb{E}^{(\Delta\theta)}\left[\mathbb{P}_{\theta,\tau_0}(\mathcal{A})\,+\,\mathbb{P}_{\theta,\tau_1}(\bar{\mathcal{A}})\right]\,-$$

$$-\,\frac{\delta_N^{-\beta}}{4}\,\mathbb{E}^{(\Delta\theta)}\left[\mathbf{1}\left\{\theta\notin\Theta^{(0)}(\varepsilon_N)\right\}\left(\mathbb{P}_{\theta,\tau_0}(\mathcal{A})\,+\,\mathbb{P}_{\theta,\tau_1}(\bar{\mathcal{A}})\right)\right]\geq$$

$$\geq \frac{1}{4}\,\mathbb{E}^{(\Delta\theta)}\left[\mathbb{P}_{\theta,\tau_0}(\mathcal{A})\,+\,\mathbb{P}_{\theta,\tau_1}(\bar{\mathcal{A}})\right]\,-\,\frac{\delta_N^{-\beta}}{2}\,\mathbb{E}^{(\Delta\theta)}\left[\mathbf{1}\left\{\theta\notin\Theta^{(0)}(\varepsilon_N)\right\}\right], \tag{4.5}$$

where we assumed without loss of generality that $0\leq\hat\tau_N(x_1)\leq 1$. The random event $\mathcal{A}=\left\{|\hat\tau_N(a)-\tau_0(a)|>\delta_N^{\beta}/2\right\}$ and $\bar{\mathcal{A}}$ is its compliment. Using (4.1) and (4.4), we find that the second term in (4.5) is vanishing

$$\delta_N^{-\beta}\,\mathbb{E}^{(\Delta\theta)}\left[\mathbf{1}\left\{\theta\notin\Theta^{(0)}(\varepsilon_N)\right\}\right]\,=\,\delta_N^{-\beta}\,\mathbb{P}^{(\Delta\theta)}\left(\theta\notin\Theta^{(0)}(\varepsilon_N)\right)$$
$$\leq\,\varepsilon_N^{-1}\,e^{-N}\,\leq\sqrt{N}\,e^{-N}\,\to 0\,,\quad\text{as}\quad N\to\infty.$$

Thus, it is enough to show that the lower limit of the expectation

$$\mathbb{E}^{(\Delta\theta)}\left[\mathbb{P}_{\theta,\tau_0}(\mathcal{A})\,+\,\mathbb{P}_{\theta,\tau_1}(\bar{\mathcal{A}})\right] \tag{4.6}$$

is bounded from below by a positive constant.

Define the two likelihood ratios and their expectations with respect to the distribution $\mathbb{P}^{(\Delta\theta)}$ of the random variables $\Delta\theta_j$. Since we have $j=1,2,\ldots,N$ channels, the likelihood ratios are

$$\Lambda_i\,=\,\frac{d\mathbb{P}_{\Delta\theta,\tau_i}}{d\mathbb{P}_{0,\tau_i}}\,=$$
$$=\,\exp\left\{\sum_{j=1}^{N}\left[\iint_{G_i}\Delta\theta_j\dot{W}_j\,dx_2\,dx_1\,-\,\frac{1}{2}\iint_{G_i}\Delta\theta^2\,(\theta,\tau)\,dx_2\,dx_1\right]\right\},\quad i=0,1.$$

The corresponding expectations with respect to the distribution $\mathbb{P}^{(\Delta\theta)}$ we denote by

$$Z_0\,=\,\mathbb{E}^{(\Delta\theta)}\,[\,\Lambda_0\,]\qquad\text{and}\qquad Z_1\,=\,\mathbb{E}^{(\Delta\theta)}\,[\,\Lambda_1\,]\,.$$

With these notations we can rewrite (4.6) as

$$\mathbb{E}^{(\Delta\theta)}\left[\mathbb{E}_{0,\tau_0}\,(\,\Lambda_0\,\mathbf{1}\,(\,\mathcal{A})\,)\,+\,\mathbb{E}_{0,\tau_1}\,(\,\Lambda_1\,\mathbf{1}\,(\,\bar{\mathcal{A}})\,)\,\right]\,=$$
$$=\mathbb{E}_{0,\tau_0}\,(\,Z_0\,\mathbf{1}\,(\mathcal{A})\,)\,+\,\mathbb{E}_{0,\tau_1}\,(\,Z_1\,\mathbf{1}\,(\bar{\mathcal{A}})\,)\,\geq$$
$$\geq\mathbb{E}_{0,\tau_0}\,(\,\min\{\,Z_0,\;Z_1\,\})\,.$$

The expectations $\mathbb{E}_{0,\tau_0}$ and $\mathbb{E}_{0\ \tau_1}$ are identical since there is no difference between the case when we have no jump at $\tau_0(x_1)$ or $\tau_1(x_1)$. That is why, it is enough to estimate from below the minimum of Z_0 and Z_1:

$$\mathbb{E}_{0,\tau_0}[\min\{Z_0, Z_1\}] = \mathbb{E}_{0,\tau_0}\left[\frac{Z_0 + Z_1}{2} - \frac{|Z_1 - Z_0|}{2}\right] =$$

$$= \mathbb{E}_{0,\tau_0}\mathbb{E}^{(\Delta\theta)}\frac{d\mathbb{P}_{\Delta\theta,\tau_0}}{d\mathbb{P}_{0,\tau_0}} - \frac{1}{2}\mathbb{E}_{0,\tau_0}|Z_1 - Z_0| =$$

$$= 1 - \frac{1}{2}\mathbb{E}_{0,\tau_0}\mathbb{E}^{(\Delta\theta)}\left[\frac{d\mathbb{P}_{\Delta\theta,\tau_0}}{d\mathbb{P}_{0,\tau_0}}\left|\frac{Z_1}{Z_0} - 1\right|\right] =$$

$$= 1 - \frac{1}{2}\mathbb{E}^{(\Delta\theta)}\mathbb{E}_{\Delta\theta,\tau_0}\left|\frac{Z_1}{Z_0} - 1\right| \geq 1 - \frac{1}{2}\sqrt{\mathbb{E}^{(\Delta\theta)}\mathbb{E}_{\Delta\theta,\tau_0}\left|\frac{Z_1}{Z_0} - 1\right|^2},$$

where the last inequality is due to the Schwarz inequality.

Since

$$\mathbb{E}^{(\Delta\theta)}\mathbb{E}_{\Delta\theta,\tau_0}\left[\frac{Z_1}{Z_0}\right] = \mathbb{E}^{(\Delta\theta)}\mathbb{E}_{\Delta\theta,\tau_0}\frac{d\mathbb{P}_{\Delta\theta,\tau_1}}{d\mathbb{P}_{0,\tau_1}}\cdot\frac{d\mathbb{P}_{0,\tau_0}}{d\mathbb{P}_{\Delta\theta,\tau_0}} = \mathbb{E}^{(\Delta\theta)}\mathbb{E}_{0,\tau_0}[Z_1] = 1,$$

then

$$\mathbb{E}^{(\Delta\theta)}\mathbb{E}_{\Delta\theta,\tau_0}\left|\frac{Z_1}{Z_0} - 1\right|^2 = \mathbb{E}^{(\Delta\theta)}\mathbb{E}_{\Delta\theta,\tau_0}\left[\frac{Z_1}{Z_0}\right]^2 - 1.$$

Thus, we have

$$\mathbb{E}_{0,\tau_0}[\min\{Z_0, Z_1\}] \geq 1 - \frac{1}{2}\sqrt{\mathbb{E}^{(\Delta\theta)}\mathbb{E}_{\Delta\theta,\tau_0}\left[\frac{Z_1}{Z_0}\right]^2 - 1}. \qquad (4.7)$$

We will now use the direct computations to calculate Z_0 and Z_1:

$$Z_i = \mathbb{E}^{(\Delta\theta)}\exp\left\{\sum_{j=1}^{N}\left[\iint_{G_i}\Delta\theta_j\dot{W}_j - \frac{1}{2}\iint_{G_i}\Delta\theta_j^2\right]\right\} =$$

$$= \mathbb{E}^{(\Delta\theta)}\exp\left\{\sum_{j=1}^{N}\left[\Delta\theta_j\iint_{G_i}\dot{W}_j - \frac{1}{2}\Delta\theta_j^2 S_0\right]\right\}, \qquad i = 0, 1.$$

First, for the sake of simplicity, we compute this expectation for one channel:

$$\mathbb{E}^{(\Delta\theta)}\exp\left\{-\frac{1}{2}\Delta\theta_j^2 S_i + \Delta\theta_j\iint_{G_i}\dot{W}_j\right\} =$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}}\exp\left\{-\frac{y^2}{2}S_i + y\iint_{G_i}\dot{W}_j\right\}\cdot\exp\left\{-\frac{y^2}{2\sigma^2}\right\}dy =$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}}\exp\left\{-\frac{y^2}{2}\left(S_i + \frac{1}{\sigma^2}\right) + y\iint_{G_i}\dot{W}_j\right\}dy =$$

$$= \exp\left\{\frac{\sigma^2\left(\iint_{G_i}\dot{W}_j\right)^2}{2(1 + S_i\sigma^2)}\right\}\frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}}\exp\left\{-\frac{1 + S_i\sigma^2}{2\sigma^2}\left[y - \frac{\sigma^2\iint_{G_i}\dot{W}_j}{1 + S_i\sigma^2}\right]^2\right\} =$$

$$= \frac{1}{\sqrt{1 + S_i\sigma^2}}\exp\left\{\frac{\sigma^2}{2(1 + S_i\sigma^2)}\left(\iint_{G_i}\dot{W}_j\right)^2\right\}, \qquad i = 0, 1.$$

Since $j = 1, \ldots, N$, then

$$Z_i = (1 + S_i \sigma^2)^{-\frac{N}{2}} \exp\left\{ \frac{\sigma^2}{2(1 + S_i \sigma^2)} \sum_{j=1}^{N} \left(\iint_{G_i} \dot{W}_j \right)^2 \right\}, \quad i = 0, 1. \quad (4.8)$$

If $\tau_0(x_1)$ is the true boundary function, then

$$\iint_{G_0} \dot{Y}_j \, dx_1 \, dx_2 = \Delta\theta_j S_0 + \sqrt{S_0} \xi_j^{(0)} \quad (4.9)$$

and since $S_1 - S_0 = r_0 \varepsilon_N^2$, then

$$\iint_{G_1} \dot{Y}_j \, dx_1 \, dx_2 = \iint_{G_0} \dot{Y}_j \, dx_1 \, dx_2 + \iint_{G_1 \setminus G_0} \dot{Y}_j \, dx_1 \, dx_2 =$$
$$= \Delta\theta_j S_0 + \sqrt{S_0} \xi_j^{(0)} + \sqrt{r_0} \, \varepsilon_N \, \xi_j^{(1)}, \quad (4.10)$$

where $\xi_j^{(0)}$, $\xi_j^{(1)}$ are independent standard normal random variables. Next we will use the following auxiliary result:

Lemma 4.1. *Let $\xi \sim \mathcal{N}(0,1)$ be a standard normal random variable. Assume that numbers α and β are such that $1 - \alpha\beta^2 > 0$. Then for any $\mu \in \mathbb{R}$*

$$\mathbb{E}\left[\exp\left\{ \frac{\alpha}{2} (\mu + \beta\xi)^2 \right\} \right] = \frac{1}{\sqrt{1 - \alpha\beta^2}} \exp\left\{ \frac{\alpha\mu^2}{2(1 - \alpha\beta^2)} \right\}.$$

Now we are ready to estimate the expectation with respect to distribution $\mathbb{P}_{\Delta\theta, \tau_0}$ on the right-side of (4.7). From (4.8)-(4.10) we find that

$$\mathbb{E}_{\Delta\theta, \tau_0} \left[\frac{Z_1}{Z_0} \right]^2 =$$
$$= \prod_{j=1}^{N} \frac{1 + S_0 \sigma^2}{1 + S_1 \sigma^2} \mathbb{E}_{\Delta\theta, \tau_0} \exp\left\{ \frac{\sigma^2}{1 + S_1 \sigma^2} \left(\iint_{G_1} \dot{W}_j \right)^2 - \frac{\sigma^2}{1 + S_0 \sigma^2} \left(\iint_{G_0} \dot{W}_j \right)^2 \right\} =$$
$$= \prod_{j=1}^{N} \frac{1 + S_0 \sigma^2}{1 + S_1 \sigma^2} \mathbb{E}_{\Delta\theta, \tau_0} \exp\left\{ - \frac{\sigma^2 \left[\mu_j^{(0)} \right]^2}{1 + S_0 \sigma^2} \right\} \cdot \exp\left\{ \frac{\sigma^2}{1 + S_1 \sigma^2} \left[\mu_j^{(0)} + \sqrt{r_0} \varepsilon_N \xi_j^{(1)} \right]^2 \right\},$$

where $\mu_j^{(0)} = \Delta\theta_j \, S_0 + \sqrt{S_0} \, \xi_j^{(0)}$. Applying Lemma (4.1) with $\alpha = \frac{2\sigma^2}{1 + S_1 \sigma^2}$, $\beta =$

$\sqrt{r_0}\varepsilon_N$ and $\mu = \mu_j^{(0)}$ and averaging over $\xi_j^{(1)}$ we get

$$\mathbb{E}_{\Delta\theta, r_0}\left[\frac{Z_1}{Z_0}\right]^2 = \left(\frac{1+S_0\sigma^2}{1+S_1\sigma^2}\right)^N \left(\frac{1+S_1\sigma^2}{1+S_1\sigma^2 - 2\sigma^2 r_0\varepsilon_N^2}\right)^{N/2} \cdot$$
$$\cdot \prod_{j=1}^N \mathbb{E}_{\Delta\theta, r_0} \exp\left\{\left(-\frac{\sigma^2}{1+S_0\sigma^2} + \frac{\sigma^2}{1+S_1\sigma^2 - 2\sigma^2 r_0\varepsilon_N^2}\right)\left(\mu_j^{(0)}\right)^2\right\} =$$
$$= (1+S_0\sigma^2)^N (1+S_1\sigma^2)^{-N/2} (1+S_1\sigma^2 - 2\sigma^2 r_0\varepsilon_N^2)^{-N/2} \cdot$$
$$\cdot \prod_{j=1}^N \mathbb{E}_{\Delta\theta, r_0} \exp\left\{\frac{\alpha_0}{2}\left(\mu_j^{(1)} + \sqrt{S_0}\xi_j^{(0)}\right)\right\} =$$
$$= \left(1 - \frac{\sigma^4 r_0^2\varepsilon_N^4}{(1+S_0\sigma^2)^2}\right)^{-N/2} \cdot \prod_{j=1}^N \mathbb{E}_{\Delta\theta, r_0} \exp\left\{\frac{\alpha_0}{2}\left(\mu_j^{(1)} + \sqrt{S_0}\xi_j^{(0)}\right)\right\},$$

where $\mu_j^{(1)} = \Delta\theta_j S_0$ and

$$\frac{\alpha_0}{2} = \frac{\sigma^2}{1+S_1\sigma^2 - 2\sigma^2 r_0\varepsilon_N^2} - \frac{\sigma^2}{1+S_0\sigma^2} = \frac{\sigma^4 r_0\varepsilon_N^2}{(1+S_0\sigma^2)(1+S_1\sigma^2 - 2\sigma^2 r_0\varepsilon_N^2)}.$$

Note that $\sigma^4 r_0\varepsilon_N^2 = r_0/N$ as $N \to \infty$, while the denominator of the latter formula approaches at 1, so $\alpha_0 \sim 2r_0/N$.

Once again, applying lemma (4.1) with $\alpha = \alpha_0$, $\beta = \sqrt{S_0}$ and $\mu = \mu_j^{(1)}$, and averaging now over $\xi_j^{(0)}$ to obtain

$$\mathbb{E}_{\Delta\theta, r_0}\left[\frac{Z_1}{Z_0}\right]^2 = \left(1 - \frac{\sigma^4 r_0^2\varepsilon_N^4}{(1+S_0\sigma^2)^2}\right)^{-N/2} \cdot (1-\alpha_0 S_0)^{-N/2} \cdot \exp\left\{\frac{\alpha_0 S_0^2\|\Delta\theta\|^2}{2(1-\alpha_0 S_0)}\right\}. \tag{4.11}$$

For N large the first factor on the right-hand side of (4.11) is equivalent to

$$\left(1 - \frac{\sigma^4 r_0^2\varepsilon_N^4}{(1+S_0\sigma^2)^2}\right)^{-N/2} \sim \exp\left\{\frac{r_0^2\varepsilon_N^2}{2}\right\} \leq \exp\left\{\frac{r_0^2}{2}\right\}.$$

The second factor in (4.11) has a finite limit:

$$(1-\alpha_0 S_0)^{-N/2} \sim \left(1 - \frac{2r_0 S_0}{N}\right)^{-N/2} \sim \exp\{r_0 S_0\} < \infty, \quad \text{as} \quad N \to \infty.$$

Using the Lemma 4.1 one more time with $\alpha = \frac{\alpha_0 S_0^2}{1-\alpha_0 S_0}$, $\mu = 0$ and $\beta = \sigma$ the last factor in (4.11) has expectation

$$\mathbb{E}^{(\Delta\theta)} \exp\left\{\frac{\alpha_0 S_0^2\|\Delta\theta\|^2}{2(1-\alpha_0 S_0)}\right\} = \left(1 - \frac{\alpha_0 S_0^2\sigma^2}{1-\alpha_0 S_0}\right)^{-N/2} \sim \left(1 - \frac{2r_0 S_0^2\sigma^2}{N}\right)^{-N/2} \sim$$
$$\sim \exp\{r_0 S_0^2\sigma^2\} \to 1 \quad \text{as } N \to \infty. \tag{4.12}$$

Choose $r_0 = \int_{-1}^{1} \varphi(x_1)\,dx_1$ so that

$$\mathbb{E}^{(\Delta\theta)}\,\mathbb{E}_{\Delta\theta, r_0}\left[\frac{Z_1}{Z_0}\right]^2 \le \exp\left\{\frac{r_0^2}{2}\right\} \cdot \exp\{r_0 S_0\}\,\mathbb{E}^{(\Delta\theta)}\,\exp\{r_0\, S_0^2\,\sigma^2\} \le 2\,.$$

Now the expectation of the minimum of Z_0 and Z_1 in (4.7) is bounded from below by a positive constant

$$\mathbb{E}_{\Delta\theta, r_0}\left[\min\ Z_1\,, Z_0\,\right] \ge C_L \ge 0.$$

Hence, the expectation in (4.6) is also bounded from below by the same constant. This completes the proof. □

5 Appendix

Proof. *(Lemma 3.1)* Add to the right-hand side of (3.2) the following sum independent of $\hat{\tau}$:

$$\frac{1}{\|\Delta\theta\|^2}\sum_{j=1}^{N}\Big(\,-\,\dot{Y}_j(x_1\,,x_2)\,[\,\theta_j''\,\mathbf{1}(x_2 < \hat{\tau}(x_1)) \,+\, \theta_j''\,\mathbf{1}(x_2 \ge \hat{\tau}(x_1))\,]+$$
$$+\frac{1}{2}\,[\,(\theta_j'')^2\,\mathbf{1}(x_2 < \hat{\tau}(x_1)) \,+\, (\theta_j'')^2\,\mathbf{1}(x_2 \ge \hat{\tau}(x_1))\,]\,\Big)\,. \qquad (5.1)$$

After that, as the easy algebra shows, the right-hand side of (3.2) turns into the right-hand side of (3.1). Next, the random field in (3.1) satisfies,

$$\dot{Z}(x_1, x_2) \;=\; \mathbf{1}(x_2 < \hat{\tau}(x_1))\,\frac{1}{\|\Delta\theta\|^2}\sum_{j=1}^{N}(\theta_j' - \theta_j'')\,[\,\theta_j'\mathbf{1}(x_2 < \tau(x_1)) \;+$$

$$+\,\theta_j''\mathbf{1}(x_2 \ge \tau(x_1)) \;-\; \frac{1}{2}(\theta_j' + \theta_j'')\,] \;+\; \mathbf{1}(x_2 < \hat{\tau}(x_1))\,\frac{1}{\|\Delta\theta\|^2}\,\sum_{j=1}^{N}(\theta_j' - \theta_j'')\,\dot{W}_j(x_1, x_2)\,.$$

Note that

$$\dot{W}(x_1\,,x_2) \;=\; \frac{1}{\|\Delta\theta\|}\,\sum_{j=1}^{N}(\theta_j' - \theta_j'')\,\dot{W}_j(x_1, x_2)$$

defines a new standard Wiener sheet, and the stochastic term equals

$$\mathbf{1}(x_2 < \hat{\tau}(x_1))\,\|\Delta\theta\|^{-1}\,\dot{W}(x_1\,,x_2)\,.$$

Finally, find that

$$\frac{1}{\|\Delta\theta\|^2}\sum_{j=1}^{N}(\theta_j' - \theta_j'')\,[\,\theta_j'\mathbf{1}(x_2 < \tau(x_1)) \;+\; \theta_j''\mathbf{1}(x_2 \ge \tau(x_1)) \;-\; \frac{1}{2}(\theta_j' + \theta_j'')\,] \;=$$

$$\frac{1}{\|\Delta\theta\|^2}\sum_{j=1}^{N}(\theta_j' - \theta_j'')^2\left[\frac{1}{2}\left(\mathbf{1}(x_2 < \tau(x_1)) - \mathbf{1}(x_2 \ge \tau(x_1))\right)\right] \;=$$

$$\frac{1}{2}\left(\mathbf{1}(x_2 < \tau(x_1)) - \mathbf{1}\{x_2 \ge \tau(x_1)\}\right)$$

and the lemma follows. □

Proof. (*Lemma 3.2*) First, note that $\dot{Z}(x_1,x_2|\hat{\tau})$ in (3.3) is computable from the data.

Put $\xi_j = \xi_j' - \xi_j'' \sim \mathcal{N}(0\,,16/t_0)$ and $\eta_j = \eta_j' + \eta_j'' \sim \mathcal{N}(0\,,16/t_0)$. Simplifying (3.3), we obtain that

$$\dot{Z}(x_1,x_2|\hat{\tau}) = \mathbf{1}(x_2 < \hat{\tau}(x_1)) \sum_{j=1}^{N} (\theta_j' + \xi_j' - \theta_j'' - \xi_j'')(\dot{Y}_j(x_1,x_2) - \frac{1}{2}(\theta_j' + \eta_j' + \theta_j'' + \eta_j'')) .$$

From (2.2), the sum on the right-hand side of the latter equation can be written as

$$\sum_{j=1}^{N} (\Delta\theta_j + \xi_j) \left[\frac{1}{2}\mathbf{1}(x_2 < \tau(x_1))(\Delta\theta_j - \eta_j) - \frac{1}{2}\mathbf{1}(x_2 \ge \tau(x_1))(\Delta\theta_j + \eta_j) + \dot{W}_j \right] =$$

$$= \frac{1}{2}\|\Delta\theta\|^2 \left[\mathbf{1}(x_2 < \tau(x_1))(1+\alpha_2) - \mathbf{1}(x_2 \ge \tau(x_1))(1+\alpha_3) \right] + \sqrt{\sum_{j=1}^{N}(\Delta\theta_j + \xi_j)^2}\, \dot{W}(x_1,x_2) ,$$

where $W(x_1,x_2)$ is a new standard Wiener sheet;

$$\alpha_2 = \frac{4}{\sqrt{t_0}} \frac{\bar{\xi} - \bar{\eta}}{\|\Delta\theta\|} - \frac{1}{\|\Delta\theta\|^2} \sum_{j=1}^{N} \xi_j\eta_j \quad \text{and} \quad \alpha_3 = \frac{4}{\sqrt{t_0}} \frac{\bar{\xi} + \bar{\eta}}{\|\Delta\theta\|} + \frac{1}{\|\Delta\theta\|^2} \sum_{j=1}^{N} \xi_j\eta_j$$

with the independent standard $(0\,,1)$-normal random variables

$$\bar{\xi} = \sqrt{t_0/16}\,\|\Delta\theta\|^{-1} \sum_{j=1}^{N} \Delta\theta_j\,\xi_j \quad \text{and} \quad \bar{\eta} = \sqrt{t_0/16}\,\|\Delta\theta\|^{-1} \sum_{j=1}^{N} \Delta\theta_j\,\eta_j .$$

Note that

$$\sum_{j=1}^{N} (\Delta\theta_j + \xi_j)^2 = \|\Delta\theta\|^2 + \frac{8}{\sqrt{t_0}}\|\Delta\theta\|\bar{\xi} + \frac{16}{t_0}(\chi_N^2 - N) + \frac{16}{t_0}N = \|\Delta\theta\|^2(1+\alpha_1) + \frac{16}{t_0}N ,$$

where

$$\alpha_1 = \frac{8\,\bar{\xi}}{\sqrt{t_0}\,\|\Delta\theta\|} + \frac{16}{t_0}\frac{\chi_N^2 - N}{\|\Delta\theta\|^2}$$

and $\chi_N^2 = (t_0/16)\sum_{j=1}^{N} \xi_j^2$ is the chi-square random variable with N degrees of freedom. To complete the proof of the lemma, we have to show that the random variables $\bar{\xi}/\|\Delta\theta\|$, $\bar{\eta}/\|\Delta\theta\|$, $\sum_{j=1}^{N} \xi_j\eta_j/\|\Delta\theta\|^2$, and $(\chi_N^2 - N)/\|\Delta\theta\|^2$ are vanishing as $N \to \infty$. From the first condition in (2.5), we have that $\|\Delta\theta\|^2 \ge \sqrt{N}$, so that the convergence to zero of the first two random variables is trivial. In what concerns the last two random variables, some calculations are necessary. Recall that $(16/t_0)\,\xi_j\,\eta_j$ is the product of two independent standard normal random variables. For whatever small $\gamma > 0$, the Chernoff bound yields the inequalities,

$$\mathbb{P}_{\theta,\tau}\Big(\frac{16}{t_0}\Big|\sum_{j=1}^{N}\xi_j\eta_j\Big| \ge \gamma\|\Delta\theta\|^2\Big) \le 2\exp\Big(-\frac{\gamma}{\sqrt{N}}\|\Delta\theta\|^2\Big)\,\mathbb{E}_{\theta,\tau}\Big[\exp\Big(\frac{16}{t_0\sqrt{N}}\sum_{j=1}^{N}\xi_j\eta_j\Big)\Big]$$

$$= 2\,\exp(-\frac{\gamma}{\sqrt{N}}\,\|\Delta\theta\|^2)\,\Big(\,\mathbb{E}_{\theta,\tau}\,[\,\exp\,(\frac{1}{2N}\,\frac{16}{t_0}\,\xi_1^2)\,]\,\Big)^N$$

$$= 2\,\exp(-\frac{\gamma}{\sqrt{N}}\,\|\Delta\theta\|^2)\,(1-\frac{1}{N})^{-N/2} \le 8\,\exp(-\frac{\gamma}{\sqrt{N}}\,\|\Delta\theta\|^2) \to 0 \text{ as } N\to\infty,$$

the latter convergence to zero being uniform over those θ's for which the convergence $\|\Delta\theta\|^2/\sqrt{N} \to \infty$ is uniform. Similarly, for all N large,

$$\begin{aligned}
&\mathbb{P}_{\theta,\tau}\,(\,|\,\chi_N^2 - N\,| \ge \gamma\|\Delta\theta\|^2\,) \le \\
&\le \exp(-N^{-1/2}\,\gamma\|\Delta\theta\|^2)\,\Big(\,\mathbb{E}_{\theta,\tau}\,[\,\exp\,(\,-\,N^{-1/2} + N^{-1/2}\,\xi_1^2\,)\,]\,\Big)^N \\
&+ \exp(-N^{-1/2}\,\gamma\|\Delta\theta\|^2)\,\Big(\,\mathbb{E}_{\theta,\tau}\,[\,\exp\,(N^{-1/2} - N^{-1/2}\,\xi_1^2\,)\,]\,\Big)^N = \\
&= \exp(-N^{-1/2}\,\gamma\|\Delta\theta\|^2)\,\Big(\,\exp\,(\,-\,N^{-1/2} - (1/2)\,\ln(1-2\,N^{-1/2})\,\Big)^N + \\
&+ \exp(-N^{-1/2}\,\gamma\|\Delta\theta\|^2)\,\Big(\,\exp\,(N^{-1/2} - (1/2)\,\ln(1+2\,N^{-1/2})\,\Big)^N = \\
&= 2\,\exp(-N^{-1/2}\,\gamma\|\Delta\theta\|^2)\,\exp\,(\,N\,(N^{-1} + o(N^{-1})\,) \le 8\,\exp(-N^{-1/2}\,\gamma\|\Delta\theta\|^2)
\end{aligned}$$

where the Maclaurence's expansion for the logarithm has been applied. This completes the proof. □

Proof. (Lemma 3.3) From Lemma 3.2 and the definition of $\mathcal{L}$, we obtain that

$$\begin{aligned}
\mathcal{L}(\hat\tau\,|\,\tau,\theta) &= \int_{a-\delta_N}^{a+\delta_N}\int_{t_0}^{1-t_0} \mathbf{1}(x_1 : \hat\tau(x_1) < \tau(x_1))\,[\,\dot Z(x_1,x_2\,|\,\hat\tau) - \dot Z(x_1,x_2\,|\,\tau)\,]\,dx_2\,dx_1 + \\
&+ \int_{a-\delta_N}^{a+\delta_N}\int_{t_0}^{1-t_0} \mathbf{1}(x_1 : \hat\tau(x_1) \ge \tau(x_1))\,[\,\dot Z(x_1,x_2\,|\,\hat\tau) - \dot Z(x_1,x_2\,|\,\tau)\,]\,dx_2\,dx_1 = \\
&= \int_{a-\delta_N}^{a+\delta_N}\int_{\hat\tau(x_1)}^{\tau(x_1)} \mathbf{1}(x_1 : \hat\tau(x_1) < \tau(x_1))\,\Big[-\frac{1}{2}\,(1+\alpha_2) + \epsilon_N\,\dot W(x_1,x_2)\Big]\,dx_2\,dx_1 + \\
&+ \int_{a-\delta_N}^{a+\delta_N}\int_{\tau(x_1)}^{\hat\tau(x_1)} \mathbf{1}(x_1 : \hat\tau(x_1) \ge \tau(x_1))\,\Big[-\frac{1}{2}\,(1+\alpha_3) + \epsilon_N\,\dot W(x_1,x_2)\Big]\,dx_2\,dx_1 = \\
&= -\,\frac{1}{2}\,d_1(\hat\tau,\tau)\,(1+\alpha_4) + \epsilon_N \int_{a-\delta_N}^{a+\delta_N} |\,\hat\tau(x_1) - \tau(x_1)|\,\dot W(x_1,x_2)\,dx_1\,,
\end{aligned}$$

where the random variable

$$\begin{aligned}
\alpha_4 = d_1^{-1}(\hat\tau\,,\,\tau)\int_{a-\delta_N}^{a+\delta_N} |\hat\tau(x_1) - \tau(x_1)|\,\big[\alpha_2\,\mathbf{1}\big(x_1 : \hat\tau(x_1) < \tau(x_1)\big) \;+ \\
+\,\alpha_3\,\mathbf{1}\big(x_1 : \hat\tau(x_1) \ge \tau(x_1)\,\big)\big]\,dx_1
\end{aligned}$$

is bounded, $|\,\alpha_4\,| \le |\,\alpha_2\,| + |\,\alpha_3\,|$. The variance of the Gaussian stochastic term equals $\epsilon_N^2\,d_1(\hat\tau\,,\,\tau)$, where ϵ_N^2 is defined in Lemma 3.2,

$$\epsilon_N^2 = \|\Delta\theta\|^{-4}\,(\,\|\Delta\theta\|^2\,(1+\alpha_1) + 16\,N/t_0\,) = \frac{16}{t_0}\,\frac{N}{\|\Delta\theta\|^4}\,(\,1 + (1+\alpha_1)\,\frac{t_0}{16}\,\frac{\|\Delta\theta\|^2}{N}\,)$$

$$= \frac{16}{t_0}\,\frac{N}{\|\Delta\theta\|^4}\,(1 + \alpha_5) \quad \text{with} \quad \alpha_5 = (1+\alpha_1)\,\frac{t_0}{16}\,\frac{\|\Delta\theta\|^2}{N} \to 0 \ \text{ as } \ N \to \infty .$$

Note that the convergence $\alpha_5 \to 0$ is uniform over those θ's for which the convergence $\|\Delta\theta\|^2/N \to 0$ in (2.5) is uniform. □

Proof. (Lemma 3.4) Put $\mu_i = \delta_N^{i-\beta}\,\tau^{(i)}(a)$ and note that $|\,\mu_i - m^{(0,i)}\,| \le 1$, $i = 0, \ldots, \beta - 1$. The Taylor's expansion of $\tau(x_1)$ can be written as

$$\tau(x_1) = \delta_N^{\beta} \sum_{i=0}^{\beta-1} \frac{\mu_i}{i!} \left(\frac{x_1 - a}{\delta_N} \right)^i + \operatorname{Rem}(x_1) \tag{5.2}$$

where the remainder term $\operatorname{Rem}(x_1)$, uniformly over $\tau \in \Sigma(\beta, L, t_0)$, satisfies

$$|\operatorname{Rem}(x_1)\,| \le \frac{L}{(\beta-1)!}\,\delta_N^{\beta}$$

for any x_1 such that $|\,x_1 - a\,| \le \delta_N$. Indeed, any function $\tau \in \Sigma(\beta)$ admits the expansion,

$$\tau(x_1) = \delta_N^{\beta} \left[\sum_{i=0}^{\beta-2} \frac{\mu_i}{i!} \left(\frac{x_1 - a}{\delta_N} \right)^i + \frac{\delta_N^{\ -1}\tau^{(\beta-1)}(x_1^*)}{(\beta-1)!} \left(\frac{x_1^* - a}{\delta_N} \right)^{\beta-1} \right]$$

with an intermediate point x_1^* located between x_1 and a. Note that

$$|\,\delta_N^{-1}\,\tau^{(\beta-1)}(x_1^*) - \mu_{\beta-1}\,| \le \delta_N^{-1}\,|\,\tau^{(\beta-1)}(x_1^*) - \tau^{(\beta-1)}(a)\,| \le L .$$

Thus, the absolute value of the remainder term in (5.2) is bounded by

$$|\operatorname{Rem}(x_1)\,| = \delta_N^{\beta}\,\frac{|\,\delta_N^{\ -1}\tau^{(\beta-1)}(x_1^*) - \mu_{\beta-1}\,|}{(\beta-1)!} \left(\frac{x_1^* - a}{\delta_N} \right)^{\beta-1} \le \frac{L}{(\beta-1)!}\,\delta_N^{\beta} .$$

Hence

$$d_1(\tau, \tau^{(0)}) = \int_{a-\delta_N}^{a+\delta_N} |\,\tau(x_1) - \tau^{(0)}(x_1)\,|\,dx_1 \le$$

$$\le \delta_N^{\beta} \int_{a-\delta_N}^{a+\delta_N} \left| \sum_{i=0}^{\beta-1} \frac{\mu_i - m^{(0,i)}}{i!} \left(\frac{x_1 - a}{\delta_N} \right)^i \right| dx_1 + \int_{a-\delta_N}^{a+\delta_N} |\operatorname{Rem}(x_1)\,|\,dx_1$$

$$\le (2\,\delta_N)\left[\beta\,\delta_N^{\beta} + \frac{L}{(\beta-1)!}\,\delta_N^{\beta}\right] = C_T\,\delta^{\beta+1} .$$

□

Proof. (Lemma 3.5) Applying the definitions of $\hat{\tau} \in \mathcal{M}_\beta$ and $\tau^{(0)}$, we obtain that

$$d_1(\hat{\tau}, \tau^{(0)}) = \delta_N^{\beta} \int_{a-\delta_N}^{a+\delta_N} \left| \sum_{i=0}^{\beta-1} \frac{m_i}{i!} \left(\frac{x_1 - a}{\delta_N} \right)^i - \sum_{i=0}^{\beta-1} \frac{m^{(0,i)}}{i!} \left(\frac{x_1 - a}{\delta_N} \right)^i \right| dx_1 =$$

$$= \delta_N^{\beta+1} \int_{-1}^{1} \left| \sum_{i=0}^{\beta-1} \frac{q_i}{i!}\,t^i \right| dt = \delta_N^{\beta+1}\,\|q\| \int_{-1}^{1} \left| \sum_{i=0}^{\beta-1} \frac{q_i}{\|q\|\,i!}\,t^i \right| dt$$

with $t = (x_1 - a)/\delta_N$, an integer $q_i = m_i - m^{(0,i)}$, $i = 0, \dots, \beta - 1$, and the norm

$$\|q\| = \sum_{i=0}^{\beta-1} \frac{|q_i|}{i!}$$

Note that the absolute values of the polynomial coefficients $q_i/(\|q\|\, i!)$ do not exceed 1 and their sum is 1. The set $\mathcal{P}$ of polynomials with such coefficients defines a compact in accordance with the Arzella-Ascolli Theorem. Hence the minimum of the latter integral over $\mathcal{P}$ is strictly positive,

$$\min_{\mathcal{P}} \int_{-1}^{1} \left| \sum_{i=0}^{\beta-1} \frac{q_i}{\|q\|\, i!}\, t^i \right| dt \ge r_0 > 0\,.$$

Indeed, if this minimum is zero, it would attain at some polynomial with non-zero coefficients what is impossible. Thus, the number of points ν_k in the layer L_k does not exceed the number of the integer solutions $(q_0, \dots, q_{\beta-1})$ of the inequality

$$\frac{1}{(\beta-1)!}\, \delta_N^{\beta+1}\, r_0\, (\,|q_0| + \cdots + |q_{\beta-1}|\,) \le \delta_N^{\beta+1}\, r_0\, \|\, q\,\| \le d_1(\hat{\tau}\,, \tau^{(0)}) \le (k+1)\, z\, \delta_N^{\beta+1}$$

or

$$|q_0| + \cdots + |q_{\beta-1}| \le \frac{(k+1)\,(\beta-1)!}{r_0}\, z$$

So, each q_i may take no more than $2\,(k+1)\,(\beta-1)!\, z\,/r_0$ values which yields the bound $\nu_k \le (\,2\,(k+1)\,(\beta-1)!\, z\,/r_0\,)^{\beta}$, and for $k \ge 1$, the lemma follows with $A = (4\,(\beta-1)!/r_0)^{\beta}$. □

Proof. *(Lemma 3.6)* From Lemma 3.4, we have that $d_1(\hat{\tau}\,, \tau^{(0)}) \le (z_0 + C_T)\delta_N^{\beta+1}$. With the same notations as in the proof of Lemma 3.5, the claim of the lemma reduces to the following: Let $\int_{-1}^{1} |\, \sum_{i=0}^{\beta-1} q_i\, t^i\,/i!\,|\, dt \le z_0 + C_T$ for a polynomial with integer coefficients q_i's. Then there exists a constant C_0 such that $|\,q_0\,| \le C_0$. But this statement is obvious because, as in Lemma 3.5, with some positive constant r_0 the inequalities hold,

$$\frac{1}{(\beta-1)!}\, r_0\, |\,q_0\,| \le \frac{1}{(\beta-1)!}\, r_0\, (\,|q_0| + \cdots + |q_{\beta-1}|\,) \le z_0 + C_T\,,$$

that is $|\,q_0\,| \le (z_0 + C_T)\,(\beta-1)!/r_0 = C_0$. □

Proof. *(Lemma 3.7)* Apply Lemma 3.2. To finish the proof, it suffices to put $\alpha_6 = \alpha_2 + \epsilon_N\, \sqrt{8/t_0}\, \mathcal{N}(0,1)$ where $\mathcal{N}(0,1)$ is a standard normal random variable. Note that under the assumptions of the lemma, $\alpha_2 \to 0$ and $\epsilon_N \to 0$ uniformly over $\theta \in \Theta_0$. □

Proof. *(Lemma 4.1)* The direct integration gives us:

$$\mathbb{E}\left[\exp\left\{\frac{\alpha}{2}(\mu+\beta\xi)^2\right\}\right] = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\exp\left\{\frac{\alpha}{2}(\mu+\beta x)^2\right\}\exp\left\{-\frac{x^2}{2}\right\}dx_1 =$$

$$= \frac{1}{\sqrt{2\pi}}\exp\left\{\frac{\alpha\mu^2}{2(1-\alpha\beta^2)}\right\}\int_{\mathbb{R}}\exp\left\{-\frac{(1-\alpha\beta^2)}{2}\left(x-\frac{\alpha\mu\beta}{1-\alpha\beta^2}\right)^2\right\}dx =$$

$$= \frac{1}{\sqrt{1-\alpha\beta^2}}\exp\left\{\frac{\alpha\mu^2}{2(1-\alpha\beta^2)}\right\}.$$

□

References

[1] D. Donoho, Wedgelets: nearly minimax estimation of edges, The Annals of Statistics, Vol. 27, No. 3, 859-897, 1999.

[2] I. Gijbels, E. Mammen, B. Park and L. Simar, On estimation of monotone and concave frontier functions, Journal of American Statistical Association, 94, 445; ABI/INFORM Global, 220-228, Mar. 1999.

[3] P. Hall, P. Qui, Nonparametric estimation of a point-spread function in multivariate problems, The Annals of Statistics, Vol. 35, No. 4, 1512-1534, 2007.

[4] P. Hall, A. Meister, A ridge-parameter approach to deconvolution, The Annals of Statistics, Vol. 35, No. 4, 1535-1558, 2007.

[5] I. A. Ibragimov, R. Z. Has'minskii, Statistical Estimation, Asymptotic theory, Spriger-Verlag, 1981.

[6] R. Z. Has'minskii, V. S. Lebedev, On the properties of parametric estimators for areas of a discontinuous image, Problems of Control and Information Theory, Vol.19 (5-6), 375-385, 1990.

[7] A.P. Korostelev, On minimax estimation of a discontinuous signal, Theory Probability and its Applications, Vol 32, No. 4, 727-730 (1987).

[8] A. Korostelev, O. Lepski, A semi-parametric change-point problem, submitted to MMS.

[9] A. P. Korostelev, L. Simar and A. B. Tsybakov, Efficient estimation of monotone boundaries, The Annals of Statistics, Vol. 23, No. 2, 476-489, 1995.

[10] A.P. Korostelev, A.B.Tsybakov, Minimax theory of image reconstruction, Springer-Verlag, 1993.

[11] E. Mammen, A. B. Tsybakov, Asymptotically minimax recovery of sets with smooth boundaries, The Annals of Statistics, Vol. 23, No. 2, 502-524, 1995.

[12] H. G. Müller, Change-points in nonparametric regression, The Annals of Statistics, Vol. 20, No. 2, 737-671, 1992.

Veera Holdai veera@math.wayne.edu

Alexander Korostelev apk@math.wayne.edu