1-1-2013

Laplace deconvolution with noisy observations

Felix Abramovich

Marianna Pensky
University of Central Florida

Yves Rozenholc

Electronic Journal of Statistics
Vol. 7 (2013) 1094–1128
ISSN: 1935-7524
DOI: 10.1214/13-EJS796

Laplace deconvolution with noisy observations*

Felix Abramovich

Tel Aviv University
Department of Statistics and Operations Research
Tel Aviv 69978, Israel
e-mail: felix@post.tau.ac.il

Marianna Pensky†

University of Central Florida
Department of Mathematics and Statistics
University of Central Florida
Orlando FL 32816-1353, USA
e-mail: marianna.pensky@ucf.edu

and

Yves Rozenholc

Université Paris Descartes,
MAP5 - UMR CNRS 8145,
45 rue des Saints-Pères,
75006 Paris - France
e-mail: yves.rozenholc@parisdescartes.fr

Abstract: In the present paper we consider Laplace deconvolution problem for discrete noisy data observed on an interval whose length T_n may increase with the sample size. Although this problem arises in a variety of applications, to the best of our knowledge, it has been given very little attention by the statistical community. Our objective is to fill the gap and provide statistical analysis of Laplace deconvolution problem with noisy discrete data. The main contribution of the paper is an explicit construction of an asymptotically rate-optimal (in the minimax sense) Laplace deconvolution estimator which is adaptive to the regularity of the unknown function. We show that the original Laplace deconvolution problem can be reduced to nonparametric estimation of a regression function and its derivatives on the interval of growing length T_n. Whereas the forms of the estimators remain standard, the choices of the parameters and the minimax convergence rates, which are expressed in terms of T_n^2/n in this case, are affected by the asymptotic growth of the length of the interval.

We derive an adaptive kernel estimator of the function of interest, and establish its asymptotic minimaxity over a range of Sobolev classes. We illustrate the theory by examples of construction of explicit expressions of Laplace deconvolution estimators. A simulation study shows that, in

*We would like to thank Alexander Goldenshluger and Oleg Lepski for fruitful discussions of the paper.

†Marianna Pensky was partially supported by National Science Foundation (NSF), grant DMS-1106564.

addition to providing asymptotic optimality as the number of observations tends to infinity, the proposed estimator demonstrates good performance in finite sample examples.

AMS 2000 subject classifications: Primary 62G05, 62G20.

Keywords and phrases: Adaptivity, kernel estimation, minimax rates, Volterra equation, Laplace convolution.

Received May 2012.

Contents

1. Introduction

1.1. Formulation and motivation

Mathematical modeling of a variety of problems in population dynamics, mathematical physics, theory of superfluidity and many others fields leads to the convolution type Volterra equation of the first kind of the form

$$q(t) = \int_0^t g(t-\tau)f(\tau)d\tau, \quad t \geq 0, \tag{1.1}$$

where $q(t)$ is the known or observed function, $g(t)$ is the known kernel and $f(t)$ is the unknown function to be solved for.

Note that the LHS of equation (1.1) is well defined for any $t \geq 0$ if functions f and g are Riemann integrable on any finite sub-interval of $[0, \infty)$. In particular, f and g do not need to be absolutely or square integrable on the nonnegative half-line. Assume the existence of their Laplace transforms $\tilde{f}(s)$ and $\tilde{g}(s)$ for all $s \geq 0$, where

$$\tilde{f}(s) = \int_0^\infty e^{-sx} f(x)dx, \quad \text{and} \quad \tilde{g}(s) = \int_0^\infty e^{-sx} g(x)dx, \quad s \geq 0. \tag{1.2}$$

In the Laplace domain, equation (1.1) becomes $\tilde{q}(s) = \tilde{g}(s)\tilde{f}(s)$ and, therefore, the problem (1.1) is also known as Laplace deconvolution problem.

In practice, however, one typically has only discrete observations of the function q in (1.1) which are available only on a finite interval and, in addition, are corrupted by noise, that leads to the following discrete noisy version of equation (1.1)

$$y(t_i) = \int_0^{t_i} g(t_i - \tau) f(\tau) d\tau + \sigma\epsilon_i, \quad i = 1, \ldots, n, \tag{1.3}$$

where $0 \leq t_1 \leq \cdots \leq t_n \leq T_n$, ϵ_i are i.i.d. $N(0,1)$ variates, σ is the known constant variance and T_n may grow with n.

Equations of the form (1.3) appear in many practical applications. Investigations in this paper have been motivated by analysis of dynamic contrast enhanced imaging data and modeling of time-resolved measurements in fluorescence spectroscopy.

Example 1. Dynamic contrast enhanced imaging data (DCE-imaging). DCE-imaging is widely used in cancer research (see, e.g., [5, 22, 23, 11, 12, 37, 39] and [3]). Such imaging procedures have great potential for tumor detection and characterization, as well as for monitoring *in vivo* the effects of treatments. DCE-imaging follows the diffusion of a bolus of a contrast agent injected into a vein. At the microscopic level, for a given unit volume voxel of interest, denote by $Y(t)$ the number of particles in the voxel at time t and by $F(t)$ the c.d.f. of a random lapse of time during which a particle sojourns in the voxel of interest. Then, $F(t)$ satisfies the following equation which can be viewed as a particular case of equation (1.3):

$$Y(t_i) = \int_0^{t_i} AIF(t - \tau)(1 - F(\tau)) d\tau + \sigma\epsilon_i, \tag{1.4}$$

where $AIF(t)$ is the Arterial Input Function which measures concentration of particles within a unit volume voxel inside a large artery and can be estimated relatively easily. Physicians are interested in a reproducible quantification of the blood flow inside the tissue which is characterized by $f(t) = 1 - F(t)$, since this quantity is independent of the number of particles of contrast agent injected into the vein.

Example 2. Time-resolved measurements in fluorescence spectroscopy. Time-resolved measurements in fluorescence spectroscopy are widely used for studies of biological macromolecules and for cellular imaging (see, e.g., [1, 2, 20, 35, 38], and also the monograph of [27] and references therein). At present, in fluorescence spectroscopy, most of the time-domain measurements are carried out using time-correlated single-photon counting. The measured intensity decay is represented by the number of photons $N(t_k)$ that were detected within the time interval $(t_k, t_k + \Delta t)$, and appears as a noisy convolution of the impulse response function $I(t)$ with a known lamp function $L(t)$

$$N(t_k) = \int_0^{t_k} L(t_k - \tau) I(\tau) d\tau + \sigma \epsilon_k.$$

The objective is to determine the impulse response function $I(x)$ that best matches the experimental data.

1.2. Difficulty of the problem

The mathematical theory of (noiseless) convolution type Volterra equations is well developed (see, e.g., [25]) and the exact solution of (1.1) can be obtained through Laplace transform. However, direct application of Laplace transform for discrete measurements faces serious conceptual and numerical problems. The inverse Laplace transform is usually found by application of tables of inverse Laplace transforms, partial fraction decomposition or series expansion (see, e.g., [41]), neither of which is applicable in the case of the discrete noisy version of Laplace deconvolution.

Formally, by extending $g(t)$ and $f(t)$ to the negative values of t by setting $f(t) = g(t) = 0$ for $t < 0$, equation (1.1) can be viewed as a particular case of the Fredholm convolution equation

$$h(t) = \int_{-\infty}^{\infty} g(t - \tau) f(\tau) d\tau, \tag{1.5}$$

whose discrete stochastic version

$$y(t_i) = \int_{-\infty}^{\infty} g(t_i - \tau) f(\tau) d\tau + \sigma \epsilon_i, \quad i = 1, \dots, n, \tag{1.6}$$

known also as Fourier deconvolution problem, has been extensively studied in the last thirty years (see, for example, [6, 9, 10, 13, 15, 16, 17, 26, 40, 42] among others; see also monograph by [36] and references therein).

Unfortunately, the existing approaches to Fourier deconvolution cannot be easily extended to solution of noisy discrete version of Laplace convolution equation (1.3). The body of work cited above addresses one of three situations: the case when functions f and g are periodic with period T, density deconvolution, and the case of random design, where t_i in (1.6) are random variables generated by some density function.

In the first setup, convolution (1.5) becomes circular convolution and measurements in equation (1.6) are taken on an interval of fixed length T, so that the problem can be solved by application of discrete Fourier transform. However, since the functions f and g are not periodic on $[0, T_n]$, the integral in the RHS of equation (1.3) is not a circular convolution and the discrete Fourier transform cannot be directly applied. Furthermore, the length of the interval T_n may grow with n that affects the convergence rates. For relatively small T_n (e.g., $T_n \sim \log n$), approximation of Fourier transform by its discrete version will be very poor which results in low convergence rates of the estimator of f.

Density deconvolution problem and nonparametric regression estimation with random measurements t_i typically assume that, as $n \to \infty$, the measurements t_i in (1.6) adequately represent the domain of $h(t)$ in (1.5). In these setups observations are absent on a particular part of the domain only if the density which generates those observations is very low. This, however, is not at all true for equation (1.3) where lack of observations for $t > T_n$ is due entirely to experimental design and has no relation to the values of the estimated function.

To the best of our knowledge, nobody tackled the problem of Fourier deconvolution (1.6) when observations t_i are non-random fixed quantities on an interval of length T_n which grows with the number of observations. In addition, we should also mention the important *causality* property of the Laplace deconvolution not shared by its Fourier counterpart, where the values of $q(t)$ for $0 \leq t \leq T_n$ depend on values of $f(t)$ for $0 \leq t \leq T_n$ only and vice versa. Finally, we show that under mild conditions, the solution of the equation (1.3) can be represented *explicitly* via derivatives of the RHS $q(t)$ that implies computational advantages of the proposed approach.

1.3. Existing results

Only few applied mathematicians took an effort to tackle the problem with discrete measurements in the LHS of (1.1). [1] applied Laplace deconvolution for the analysis of fluorescence curves and used a parametric presentation of the solution f as a sum of exponential functions with parameters evaluated by minimizing discrepancy with the RHS. In a somewhat similar manner, [34] proposed to expand the unknown solution over a wavelet basis and find the coefficients via the least squares algorithm. [33], following [44], studied numerical inversion of the Laplace transform using Laguerre functions. Finally, [28] and [7] used discretization of the equation (1.1) and applied various versions of the Tikhonov regularization technique. However, in all of the above papers, the noise in the measurements was either ignored or treated as deterministic. The presence of random noise in (1.3) makes the problem even more challenging.

Unlike Fourier deconvolution that has been intensively studied in statistical literature (see references above), Laplace deconvolution received virtually no attention within statistical framework. To the best of our knowledge, the only paper which tackles the problem is [14] which considers a noisy version of Laplace deconvolution with a very specific kernel of the form $g(t) = be^{-at}$. The authors use the fact that, in this case, the solution of the equation (1.1) satisfies a particular linear differential equation and, hence, can be recovered using $q(t)$ and its derivative $q'(t)$. For this particular kind of kernel, the authors derived convergence rates for the quadratic risk of the proposed estimators, as n increases, under the assumption that the m-th derivative of f is continuous on

$(0, \infty)$. However, they assume that data is available on the whole nonnegative half-line (i.e. $T_n = \infty$) and that m is known (i.e., the estimator is not adaptive).

1.4. Objectives and organization of the paper

For the reasons listed above, estimation of f from discrete noisy observations y in (1.3) requires development of a novel approach. The objective of the present paper is to fill the gap and to develop general statistical methodology for Laplace deconvolution problem which allows to circumvent lack of observations for $t > T_n$ and leads to effective representation of f on the interval $(0, T_n)$, no matter what value T_n takes. We establish minimax convergence rates for Laplace deconvolution setup over Sobolev classes and derive the adaptive estimator of f which is rate-optimal over the entire range of Sobolev classes. The proposed estimator is based on estimating q and its derivatives from noisy data y in (1.3), where $q(t) = (f * g)(t) = \int_0^t g(t-\tau) f(\tau) d\tau$ is the convolution of f and g. Thus, one can use the numerous existing techniques for nonparametric estimation of a function and its derivatives. In particular, we employ kernel estimators with the global bandwidth adaptively selected by Lepski procedure.

An attractive feature of the estimation technique proposed in this paper is that estimator of f is expressed *explicitly* via q and its derivatives. Another interesting aspect of the considered model (1.3) is that the data is observed on the interval of asymptotically increasing length, where $T_n \to \infty$ as $n \to \infty$. This is indeed a reasonable assumption since, as n is growing, demands on the improvements of the estimation precision require to decrease the bias by sampling $q(t)$ for larger and larger values of t. Dependence of T on n may not significantly affect estimation procedures but evidently leads to different convergence rates that are formulated in terms of T_n^2/n.

The rest of the paper is organized as follows. Section 2 delivers main results of the paper. In particular, Section 2.1 introduces notations and assumptions used throughout the paper. In Section 2.2 we derive the lower bounds on the minimax risk of estimating f in (1.3). Section 2.3 reviews some mathematical results for noiseless Laplace deconvolution relevant for constructing the proposed estimator. Section 2.4 is dedicated to explicit derivation of Laplace deconvolution estimator in model (1.3), while Section 2.5 establishes its asymptotic adaptive minimaxity over the entire range of Sobolev classes. Section 3.1 contains examples of explicit estimators of Laplace deconvolution for various types of kernels g. The results of a simulation study are presented in Section 3.2. Section 4 concludes the paper with discussion. All the proofs are given in Appendix.

2. Main results

2.1. Notations and assumptions

In this section we introduce notations and assumptions used throughout the paper.

The $L_k(\mathbb{R}^+)$-norm of the function h is denoted by $\|h\|_k$ and $\|h\|_\infty$ is the supremum norm of h. If $k = 2$ and there is no ambiguity, we shall omit the subscript in the notation of the norm, i.e. $\|h\| = \|h\|_2$. We use the standard notation $W^{r,p}(\mathbb{R}^+)$ for a Sobolev space of functions on $[0,\infty)$ that have r weak derivatives with finite L_p-norms and omit p in this notation if $p = 2$, that is, $W^r(\mathbb{R}^+) = W^{r,2}(\mathbb{R}^+)$. In addition, we shall omit $\mathbb{R}^+$ in the notations of the norms and functional spaces and, unless the opposite is stated, assume that all functions are defined on the nonnegative part of the real line.

Let $r \geq 1$ be such that

$$g^{(j)}(0) = \begin{cases} 0, & \text{if } \ j = 0, \dots, r-2, \\ B_r \neq 0, & \text{if } \ j = r-1, \end{cases} \tag{2.1}$$

with obvious modification $g(0) = B_1 \neq 0$ for $r = 1$.

Assume now the following conditions on the unknown f and the known kernel g in (1.1):

(A1) $g \in W^{r,1} \cap W^\nu$, $\nu \geq r$.

(A2) Let Ω be the collection of distinct zeros of the Laplace transform $\tilde{g}$ of g. Then all zeros s's of $\tilde{g}$ have negative real parts, i.e.,

$$s^* = \max_{s\in\Omega} Re(s) < 0.$$

(A3) $f \in W^m$ where $m \leq \nu + 1 - r$.

Finally, we impose the following assumption on T_n and design points t_i, $i = 1, \dots, n$:

(A4) Let T_n be such that $T_n \to \infty$ but $n^{-1} T_n^2 \to 0$ as $n \to \infty$ and there exist $1 \leq \mu < \infty$ such that $\max_i |t_i - t_{i-1}| \leq \mu n^{-1} T_n$.

In what follows, we use the symbol C for a generic positive constant, independent of the sample size n, which may take different values at different places.

2.2. Lower bounds for the minimax risk

In order to establish a benchmark for an estimator of an unknown function f from its noisy Laplace convolution (1.3) we derive the asymptotic minimax lower bounds for the $L_2([0,T_n])$-risk over a Sobolev ball W_A^m of radius A. It turns out that, unlike in the density deconvolution problem or Fourier deconvolution setup, the rates of convergence depend on the length of the interval T_n and are expressed in terms of the ratio T_n^2/n:

Theorem 1. *Let condition (2.1) and Assumptions (A1)–(A4) hold. Then, there exists a constant $C > 0$ such that*

$$\inf_{\hat{f}_n} \sup_{f\in W_A^m} E||\hat{f}_n - f||^2_{L_2([0,T_n])} \geq C \left(\frac{T_n^2}{n}\right)^{\frac{2m}{2(m+r)+1}}, \tag{2.2}$$

where the infimum is taken over all possible estimators $\hat{f}_n$ of f, and, therefore,

$$\inf_{\hat{f}_n} \sup_{f \in W_A^m} E||\hat{f}_n - f||^2_{L_2([0,\infty))} \geq C \left(\frac{T_n^2}{n} \right)^{\frac{2m}{2(m+r)+1}}.$$

2.3. Solution of noiseless Volterra equation

As we have already mentioned, unlike Fourier deconvolution, an estimator $\hat{f}_n$ of the unknown f in (1.3) can be obtained explicitly in the closed form. To understand the motivation for the proposed $\hat{f}_n$ we find first the exact solution of the noiseless Volterra equation (1.1).

Taking derivatives of both sides of (1.1) under (2.1) and Assumptions (A1), (A3), one obtains

$$\begin{aligned} q^{(j)}(t) &= \int_0^t g^{(j)}(t-\tau) f(\tau) d\tau, \quad j = 1, \ldots, r-1; \\ &\cdots \\ q^{(r)}(t) &= B_r f(t) + \int_0^t g^{(r)}(t-\tau) f(\tau) d\tau, \end{aligned} \tag{2.3}$$

which is the Volterra equation of the second kind. Taking higher-order derivatives, (2.3) yields

$$\begin{aligned} q^{(r+1)}(t) &= B_r f'(t) + g^{(r)}(t) f(0) + \int_0^t g^{(r)}(t-\tau) f'(\tau) d\tau, \\ &\cdots \\ q^{(r+m)}(t) &= B_r f^{(m)}(t) + \sum_{j=0}^{m-1} g^{(r+j)}(t) f^{(j)}(0) + \int_0^t g^{(r)}(t-\tau) f^{(m)}(\tau) d\tau. \end{aligned}$$

Then, under Assumptions (A1) and (A3), one has $q^{(r+m)} \in L_2$ and, hence, $q \in W^{r+m}$.

In addition, due to Assumptions (A1) and (A3), condition (2.3) implies that $q^{(r)} \in L_1$ and, therefore, one can use the following known facts from the theory of Volterra equations of the second kind:

1. there exists a unique solution ϕ of the equation

$$g^{(r)}(t) = B_r \phi(t) + \int_0^t g^{(r)}(t-\tau) \phi(\tau) d\tau \tag{2.4}$$

called a *resolvent* of $g^{(r)}$ (see Theorem 3.1 of [25]);

2. there exists a unique solution f of (2.3) and, therefore, of the original equation (1.1), which can be written as

$$f(t) = B_r^{-1} q^{(r)}(t) - B_r^{-1} \int_0^t q^{(r)}(t-\tau) \phi(\tau) d\tau \tag{2.5}$$

(see Theorem 3.5 of [25]).

Remark 1. Assumption (A2) ensures that the solution $f(t)$ of the noiseless Laplace convolution equation (1.1) is numerically stable. By Half-Line Paley-Wiener theorem (see Theorem 2.4.1 of [25]), the resolvent $\phi(\tau)$ of g in (2.4) is absolutely integrable if and only if Assumption (A2) is satisfied. If $\tilde{g}$ has roots with positive real parts, then, by Corollary 2.4.2 from the same book, $\phi(\tau)$ is growing at an exponential rate, so that $e^{-s\tau}\phi(\tau)$ is absolutely integrable for any $s > s^*$, where s^* is defined in assumption (A2).

It follows from the above that, in order to solve the noiseless Volterra equation (1.1), one only needs to determine a resolvent ϕ in (2.4) defined entirely by the r-th derivative $g^{(r)}$ of the (*known*) kernel g. Taking Laplace transform of both sides of (2.4) yields

$$\widetilde{g^{(r)}}(s) = B_r\tilde{\phi}(s) + \widetilde{g^{(r)}}(s)\tilde{\phi}(s)$$

where, due to (2.1), one has $\widetilde{g^{(r)}}(s) = s^r\tilde{g}(s) - B_r$. Therefore, $\phi(t)$ can be obtained as an inverse Laplace transform of $\tilde{\phi}$, where

$$\tilde{\phi}(s) = \frac{s^r\tilde{g}(s) - B_r}{s^r\tilde{g}(s)}. \tag{2.6}$$

Behavior of the resolvent function ϕ is thus determined by the properties $\tilde{g}$. It turns out (see, e.g., [25], Chapter 7) that, under Assumption (A2) and (2.1), $\tilde{g}$ is analytic and, hence, all its zeros are well separated. Moreover, ϕ can be presented as the sum of a polynomial of degree $(r-1)$ and an absolutely integrable function. In a variety of practical applications, the kernel g is represented by a combination of some elementary functions and, hence, $\tilde{g}$ is not an oscillating function. Hence, the number of zeros of $\tilde{g}$ is finite and, since $\tilde{g}$ is an analytic function, these zeros are of finite orders. In this case, solution f can be written explicitly as it follows from the following theorem:

Theorem 2. *Let condition (2.1) and Assumptions (A1)–(A3) hold. Then, the resolvent ϕ in (2.6) is of the form*

$$\phi(t) = \sum_{j=0}^{r-1} \frac{a_{0,j}}{j!}t^j + \phi_1(t), \tag{2.7}$$

where $\phi_1 \in L_1$. Hence, by (2.5), f in (1.1) can be recovered as

$$f(t) = B_r^{-1}\left(q^{(r)}(t) - \sum_{j=0}^{r-1} a_{0,r-1-j}q^{(j)}(t) - \int_0^t q^{(r)}(t-\tau)\phi_1(\tau)d\tau\right). \tag{2.8}$$

If, in addition, $\tilde{g}$ has a finite number M of distinct zeros s_l of orders α_l, respectively, $l = 1, \ldots, M$, then f is of the form

$$f(t) = B_r^{-1}\left(q^{(r)}(t) - \sum_{j=0}^{r-1} b_j q^{(r-1-j)}(t) - \int_0^t q(t-x)\phi_1^{(r)}(x)dx\right), \tag{2.9}$$

where $s_0 = 0$, $\alpha_0 = r$ *and*

$$\phi_1(x) = \sum_{l=1}^{M} \sum_{j=0}^{\alpha_l - 1} \frac{a_{l,j} x^j e^{s_l x}}{j!}, \tag{2.10}$$

$$a_{l,j} = \frac{1}{(\alpha_l - 1 - j)!} \frac{d^{\alpha_l - j - 1}}{ds^{\alpha_l - j - 1}} \left[(s - s_l)^{\alpha_l} \tilde{\phi}(s) \right] \Bigg|_{s = s_l}, \tag{2.11}$$

$$b_j = a_{0,j} + \sum_{l=1}^{M} \sum_{i=0}^{\min(j, \alpha_l - 1)} \binom{j}{i} a_{l,i} s_l^{j-i}. \tag{2.12}$$

Remark 2. Note that in Theorem 2, Assumption (A1) and condition (2.1) are essential for explicit construction of estimators. However, calculations in (2.7)–(2.12) can be carried out without Assumption (A2) being valid. Assumption (A2) is only needed to ensure that $\phi_1 \in L_1$. In particular, if the number of zeros is finite, then $Re(s_l) < 0$, $l = 1, \ldots, M$, implies that ϕ_1 in (2.10) is a sum of products of polynomials and exponentials with powers having negative real parts and, hence, $\phi_1 \in L_1 \cap L_2$. If some of zeros have positive real parts, expansions (2.11) and (2.12) in Theorem 2 will still be valid but $\phi_1^{(r)}$ will contain exponential terms with positive powers that will grow and magnify the errors of estimating q as t tends to infinity.

2.4. Adaptive estimation of Laplace deconvolution

Theorem 2 leads to an estimator $\hat{f}_n$ in (1.3) of the semi-explicit form

$$\hat{f}_n(t) = B_r^{-1} \left(\widehat{q^{(r)}}(t) - \sum_{j=0}^{r-1} a_{0,r-1-j} \widehat{q^{(j)}}(t) - \int_0^t \widehat{q^{(r)}}(t - \tau) \phi_1(\tau) d\tau \right), \tag{2.13}$$

where $\widehat{q^{(j)}}(t)$ are some estimators of $q^{(j)}(t)$, $j = 0, \ldots, r$, and the function ϕ_1 is expressed in terms of the inverse Laplace transform of the completely known function $\tilde{\phi}$ defined in (2.6). Under the additional (usually satisfied) condition that $\tilde{g}$ has a finite number of zeros, the second statement of Theorem 2 leads to an explicit expression for the estimator with ϕ_1 defined by (2.10):

$$\hat{f}_n(t) = B_r^{-1} \left(\widehat{q^{(r)}}(t) - \sum_{j=0}^{r-1} b_j \widehat{q^{(r-1-j)}}(t) - \int_0^t \widehat{q}(t - x) \phi_1^{(r)}(x) dx. \right) \tag{2.14}$$

Note that, unlike (2.13), the integral term in (2.14) involves q rather than $q^{(r)}$ and, hence, the boundary effects of estimating derivatives do not propagate to interior points of the interval $[0, T_n]$.

Laplace deconvolution can be therefore reduced to nonparametric estimation of $q = f * g \in W^{r+m}$ (see Section 2.3) and its derivatives of orders up to r from the discrete noisy data in the model

$$y(t_i) = q(t_i) + \sigma \epsilon_i, \quad i = 1, \ldots, n,$$

where $0 \le t_1 \le \cdots \le t_n \le T_n$, ϵ_i are i.i.d. $N(0,1)$ variates and $\sigma > 0$ is known. This is a well-studied problem, and estimation can be carried out by a number of various approaches, e.g., kernel estimation, splines, local polynomials, wavelets, etc.

It is important to note however that for the problem at hand, the data is sampled on an interval of asymptotically increasing length that calls for necessary modifications of traditional estimators and affects their global convergence rates on the interval $[0, T_n]$ which are expressed in terms of $n^{-1}T_n^2$.

For illustration, we consider kernel estimation with the global bandwidth selected adaptively by Lepski technique. To estimate the j-th derivative of $q(t)$, $j = 0, \ldots, r$, $t \in [0, T_n]$, choose a kernel function K_j (not to be confused with the convolution kernel g) of order (L, j) with $L > r$ satisfying the following conditions:

(K1) supp $(K_j) = [-1, 1]$, K_j is twice continuously differentiable and $\int K_j^2(t)dt < \infty$.

(K2) $\int t^l K_j(t)dt = \begin{cases} 0, & l = 0, \ldots, j-1, j+1, \ldots, L-1, \\ (-1)^j j!, & l = j. \end{cases}$

Construction of such kernels is described in, e.g., [19].

Define a well-known Priestley-Chao type kernel estimator of $q^{(j)}$ with a (global) bandwidth λ_j:

$$\widehat{q_\lambda^{(j)}}(t) = \frac{1}{\lambda_j^{j+1}} \sum_{i=1}^{n} K_j\left(\frac{t - t_i}{\lambda_j}\right)(t_i - t_{i-1})y_i. \tag{2.15}$$

Certain routine boundary corrections are required for t close to the boundaries (see [18] for details).

We utilize a general methodology developed by Lepski (e.g., [31]) for data-driven selection of a bandwidth λ_j in (2.15). In particular, we apply the global bandwidth version of [32] procedure and modify it also for estimating derivatives.

The resulting procedure for choosing λ_j in (2.15) can be described as follows. For each j, $0 \le j \le r$, and the corresponding kernel K_j of order (L, j), $L > r$, consider the geometric grid of bandwidths Λ_j, where

$$\Lambda_j = \{\lambda_l = a^{-l},\ l = 0, 1, \ldots, J_n;\ J_n = (2j+1)^{-1} \log_a(n\,\sigma^{-2}\,T_n^{-2})\}, \tag{2.16}$$

and $a > 1$ is an arbitrary constant. Smaller values of a allow a finer choice of the optimal bandwidth but increase computational complexity. Note that cardinality of Λ_j does not exceed $\log_a n$ since card$(\Lambda_j) = 1 + J_n \le \log_a n$. Define

$$\hat{\lambda}_{j,n} = \max\left\{\lambda \in \Lambda_j : \|\widehat{q_\lambda^{(j)}} - \widehat{q_h^{(j)}}\|^2_{[0,T_n]} \le \frac{4\,C_j^2\,\sigma^2 T_n^2}{n\,h^{2j+1}} \text{ for all } h \in \Lambda_j, \left(\frac{\sigma^2 T_n^2}{n}\right)^{\frac{1}{2j+1}} \le h < \lambda\right\}, \tag{2.17}$$

where constants C_j are such that

$$C_j^2 > \mu^2 \|K_j\|^2 \tag{2.18}$$

and μ is defined in Assumption (A4).

We then estimate $q^{(j)}$ by

$$\widehat{q^{(j)}_{\hat{\lambda}_{j,n}}}(t) = \frac{1}{\hat{\lambda}_{j,n}^{j+1}} \sum_{i=1}^{n} K_j \left(\frac{t - t_i}{\hat{\lambda}_{j,n}} \right) (t_i - t_{i-1}) y_i, \quad l = 0, \dots, r, \tag{2.19}$$

and plug (2.19) into (2.13) or (2.14).

Note that the resulting estimators $\hat{f}_n$ are inherently adaptive to the smoothness of the underlying function f in (1.3) which is rarely known in practice.

2.5. Adaptive minimaxity

The following theorem establishes the upper bound for the $L_2([0, T_n])$-risk of the estimator $\hat{f}_n$ defined in Section 2.4 over Sobolev classes:

Theorem 3. *Let condition (2.1) and Assumptions (A1)-(A4) hold. Consider kernels K_j, $j = 0, \dots, r$ of orders (L, j), $L > r$ satisfying the conditions (K1) and (K2). Let $\hat{f}_n$ be the estimator of f of the form (2.13) or (2.14), where $\widehat{q^{(j)}}(t)$'s are given by (2.19). Then, for all $1 \le m \le \min(L, \nu + 1) - r$, and $A > 0$, one has*

$$\sup_{f \in W_A^m} E\|\hat{f}_n - f\|^2_{L_2([0,T_n])} = O\left(\left(\frac{T_n^2}{n} \right)^{\frac{2m}{2(m+r)+1}} \right). \tag{2.20}$$

Under the additional conditions on f and T_n, the results of Theorem 3 can be easily extended to the entire nonnegative half-line:

Corollary 1. *Let conditions of Theorem 3 hold and also there exists $\rho \ge 1$ such that $\int_0^\infty t^{2\rho} f^2(t) dt < \infty$ and $\lim_{n\to\infty} T_n^{-2\rho} n < \infty$. Let $\hat{f}_n$ be as in Theorem 3 for $t \le T_n$ and $\hat{f}_n \equiv 0$ for $t > T_n$. Then,*

$$\sup_{f \in W_A^m} E\|\hat{f}_n - f\|^2_{L_2([0,\infty))} = O\left(\left(\frac{T_n^2}{n} \right)^{\frac{2m}{2(m+r)+1}} \right)$$

for all $1 \le m \le \min(L, \nu + 1) - r$ and $A > 0$.

Note that the upper bounds established in Theorem 3 and Corollary 1 coincide with the minimax lower bound for the risk obtained in Theorem 1 and, thus, cannot be improved. Hence, the derived Laplace deconvolution estimators are asymptotically adaptively minimax over the entire range of Sobolev classes.

3. Examples and simulation study

3.1. Examples of explicit Laplace deconvolution estimators

In what follows, we shall consider two examples of construction of explicit estimators of f in the Laplace convolution problem.

Example 1. Consider (1.3) with

$$g(t) = (bt - \sin(bt))e^{-at}, \quad a > 0.$$

It is easy to see that $r = 4$ and $B_4 = b^3$ in (2.1), and $\tilde{g}$ is of the form

$$\tilde{g}(s) = b^3(s+a)^{-2}((s+a)^2 + b^2)^{-1}. \tag{3.1}$$

Hence, $\tilde{g}(s)$ has no zeros and one can use Theorem 2 for recovering and estimating f. By (2.6) one has

$$\widetilde{\phi}(s) = -\left(\frac{4a}{s} + \frac{6a^2+b^2}{s^2} + \frac{4a^3+2ab^2}{s^3} + \frac{a^4+a^2b^2}{s^4}\right),$$

so that, in (2.9) and (2.14), one has $\alpha_{0,0} = -4a$, $\alpha_{0,1} = -(6a^2+b^2)$, $\alpha_{0,2} = -(4a^3+2ab^2)$, $\alpha_{0,3} = -(a^4+a^2b^2)$ and $\phi_1(x) = 0$. Hence, using (2.14)

$$\begin{aligned} \hat{f}_n(t) &= b^{-3}\Big[\widehat{q^{(4)}}_{\hat{\lambda}_{n,4}}(t) + 4a\widehat{q'''}_{\hat{\lambda}_{n,3}}(t) + (6a^2+b^2)\widehat{q''}_{\hat{\lambda}_{n,2}}(t) \\ &\quad + (4a^3+2ab^2)\widehat{q'}_{\hat{\lambda}_{n,1}}(t) + (a^4+a^2b^2)\widehat{q}_{\hat{\lambda}_{n,0}}(t)\Big], \end{aligned}$$

where $\hat{\lambda}_{n,l}$, $l = 0, 1, \ldots, 4$, are defined in (2.17). The rate of convergence of $\hat{f}_n$ over W^m is given by (2.20) with $r = 4$ and is $O\left(\left(\frac{T_n^2}{n}\right)^{\frac{2m}{2m+9}}\right)$.

Example 2. Consider (1.3) with

$$g(t) = e^{-at}t^{r-1}\sum_{j=0}^{k} \frac{\rho_j}{(j+r-1)!}t^j, \quad a > 0, \tag{3.2}$$

where $k \geq 0$ and $r \geq 1$ are integers and $\rho_0 = 1$. In this case, (2.1) holds with $B_r = 1$ and

$$\tilde{g}(s) = (s+a)^{-(k+r)}\mathcal{P}(s),$$

where

$$\mathcal{P}(s) = \sum_{j=0}^{k} \rho_j (s+a)^{k-j}. \tag{3.3}$$

Therefore,

$$\widetilde{\phi}(s) = \frac{s^r\mathcal{P}(s) - (s+a)^{k+r}}{s^r\mathcal{P}(s)},$$

In particular, for $k=0$ and $r=1$, $\mathcal{P}(s)$ has no roots, so that $b_0=-a$ and we recover the result of [14]: $f(x)=q'(t)+aq(t)$. For $k=1$, $\rho_0=1$ and $\rho_1=b$, one has $g(t)=e^{-at}(bt+1)$ and $\tilde{g}(s)=(s+a)^{-2}(s+a+b)$, so that $\mathcal{P}(s)$ has a single root $s_1=-(a+b)$ of multiplicity $\alpha_1=1$. Hence, $b_0=b^2(a+b)$, $a_{1,0}=-b^2$ and $\phi_1(x)=-b^2(a+b)^{-1}e^{-(a+b)x}$ in formula (2.9) leading to the estimator of f of the form

$$\hat{f}_n(t)=\widehat{q'}_{\hat{\lambda}_{n,1}}(t)+(a-b)\widehat{q}_{\hat{\lambda}_{n,0}}(t)+b^2\int_0^t \widehat{q}_{\hat{\lambda}_{n,0}}(t-x)e^{-(a+b)x}dx. \tag{3.4}$$

The asymptotic minimax rate of convergence of $\hat{f}_n$ in (3.4) over W^m is $O\big((n^{-1}T_n^2)^{\frac{2m}{2m+3}}\big)$.

For general values of k and r, the exact form of the solution (2.9) strongly depends on the roots of the polynomial $\mathcal{P}(s)$ given by (3.3). Assume that $\mathcal{P}(s)$ has k distinct roots. Then, $\mathcal{P}(s)=\prod_{l=1}^k(s-s_l)$ and $1/\tilde{g}(s)$ allows a partial fraction decomposition

$$\frac{1}{\tilde{g}(s)}=\sum_{j=0}^r \alpha_j(s+a)^j+\sum_{l=1}^k\frac{\beta_l}{s-s_l}. \tag{3.5}$$

By observing that $\sum_{j=0}^r\alpha_j s^j$ is the quotient of s^{r+k} and $\sum_{j=0}^k\rho_j s^{k-j}$, one can recursively evaluate α_j, $j=1,\dots,r$, in (3.5) as

$$\alpha_r=1,\quad \alpha_{r-l}=-\sum_{j=\max(0,l-k)}^{l-1}\alpha_{r-j}\rho_{l-j},\quad l=1,\dots,r.$$

The values of β_l can be obtained by multiplying both sides of equation (3.5) by $\mathcal{P}(s)/(s-s_l)$ and setting $s=s_l$:

$$\beta_l=(s_l+a)^{k+r}\prod_{\substack{j=1\\ j\neq l}}^k(s_l-s_j)^{-1},\quad l=1,\dots,k.$$

The respective expression for f is of the form $f=f_1+f_2$ where

$$f_1(t) = q^{(r)}(t)+\sum_{l=0}^{r-1}q^{(l)}(t)\sum_{j=l}^r\binom{j}{l}a^{j-l}\alpha_j, \tag{3.6}$$

$$f_2(t) = \sum_{l=1}^k\beta_l\int_0^t e^{s_l x}q(t-x)dx, \tag{3.7}$$

which can easily be reduced to representation (2.9).

Under assumptions (A2) and (A3), the asymptotic minimax rate of convergence of the estimator (3.4) is provided by Theorem 3.

3.2. Simulation study

In this section we present the results of a simulation study to illustrate finite sample performance of the Laplace deconvolution procedure developed above.

First, we consider the data simulated according to the model (1.3) with five convolution kernels $g_1, \ldots, g_5$, where

$$g_1(t) = e^{-5t}(2t - \sin(2t)), \qquad g_2(t) = e^{-5t}, \qquad g_3(t) = e^{-t}(2t+1).$$

Kernel g_1 mimics an ideal behaviour of AIF in the DCE-imaging (see Example 1 in Section 1.1), while g_2 and g_3 are examples of kernels considered in Example 2 from Section 3.1. In particular, g_2 corresponds to the framework of [14]. Kernels g_4 and g_5 also fall within the general form of Example 2 from Section 3.1 with $r = 3$ and were defined by the k roots $s_1, \ldots, s_k$ of the polynomial $\mathcal{P}(s)$ in (3.3). For g_4 we considered four roots $(-4 \pm 2.5i, -0.75 \pm 1.5i)$, while for g_5 we added two more conjugate roots $-2 \pm 2i$. Both g_4 and g_5 can be seen as more realistic scenarios in the DCE-imaging. All the five kernels are presented on Figure 1.

The chosen true functions f in (1.3) are $f_1(t) = t^2 e^{-t}$, $f_2(t) = 1 - \Gamma_{2,2}(t)$ and $f_3(t) = 1 - \Gamma_{3,0.75}(t)$, where $\Gamma_{\alpha,\theta}$ is the c.d.f of the Gamma distribution with the shape parameter α and the scale parameter θ (see Figure 2). Functions f_2 and f_3 mimic sojourn time distributions of the particles of a contrast agent in DCE-imaging experiments, while f_1 is aimed to be a more general case.

The Laplace convolution $q = f * g$ which produces observations in (1.3) has been numerically computed using trapezoidal rule for approximation of the integral. The noise levels for each of the kernels $g_1, \ldots, g_5$ was chosen as $\sigma_0(g_j)/2^i$, $i = 0, \ldots 4$; $j = 1, \ldots, 5$, where the nominal noise levels $\sigma_0(g_j)$ were 0.001, 0.1, 0.01, 0.002, 0.002 for $g_1, \ldots, g_5$ respectively. We ran simulations with $n = 100$ and $n = 250$ and regular design for the t_i equally spaced between 0 and $T_n = 10$.

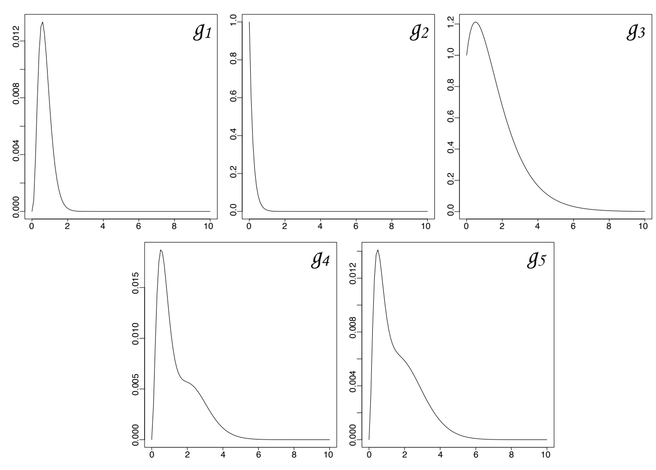

FIG 1. *From left to right and up to down: the kernels g_1 to g_5.*

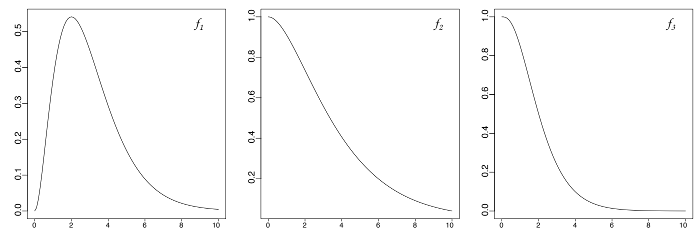

FIG 2. *The true unknown function f from left to right: f_1 to f_3.*

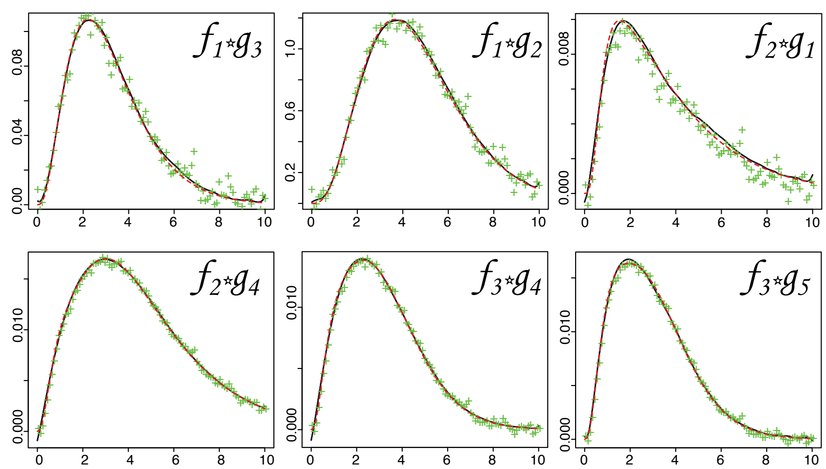

FIG 3. *Laplace convolution q (dotted red line) of the known kernel g and the unknown function f, $n = 100$ noisy observations of q (green pluses) and estimated value of q (black line). The choice of f and g used for each simulation are specified on each sub-figure by the convolution product. The noise level has been chosen as $\sigma_0(g_j)/2$ for the three left figures and $\sigma_0(g_j)/8$ for the three right figures.*

Following construction in [19], we derived kernels K_j of orders (L, j) for estimating the derivatives $q^{(j)}$ of q, $j = 0, \dots, r$ for various values of L. In our simulations we used $L = 8$ as an upper bound of the regularity of the kernel since higher values of L lead to numerically unstable computations and/or provide very little advantage in terms of precision. Finally, we used boundary kernels in order to stabilized the computations as suggested in [19]. In all simulations, due to the regular fixed design, $\mu = 1$ in Assumption (A4). We chose $a = 1.2$ in (2.16) and $C_j = 1$ in (2.18). Since the constant 4 in the Lepski's threshold in (2.17) is known to be too large for practical applications, we tried several values and "tuned" it to 3.

Figures 3 and 4 provide examples of deconvolution estimators based on single samples. Figure 5 illustrates that deconvolution estimators show good precision

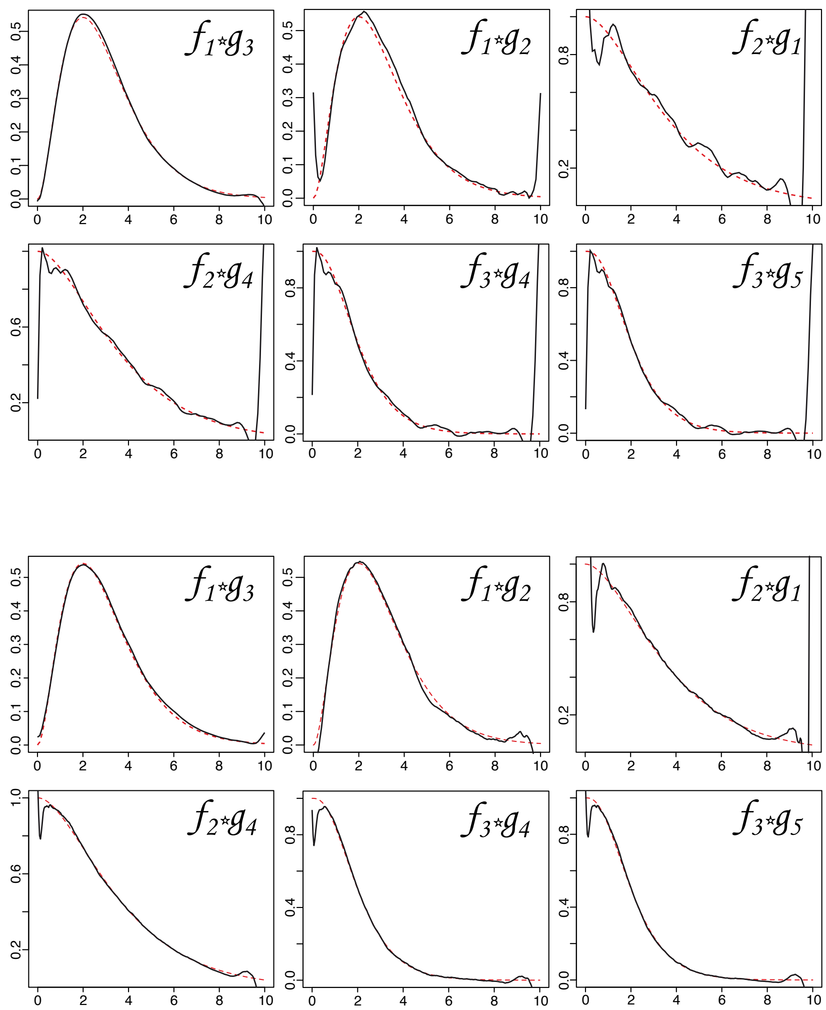

Fig 4. *True unknown f (red dotted line) and its estimate (plain black line) for $n = 100$ (top two lines) and $n = 250$ (bottom two lines).*

although boundary effects in estimating high-order derivatives remain despite the use of boundary kernels.

For each combination of true function f, kernel g, sample size n and the noise level, we ran 400 simulations and calculated mean square errors. In order to remove the influence of boundary effects (see comments above), we did not include 20% of the boundary points (10% at each boundary). The box-plots of the resulting mean square errors are presented on Figure 6. Table 1 shows the average mean square errors and standard deviations (in parentheses) over 400 simulation runs.

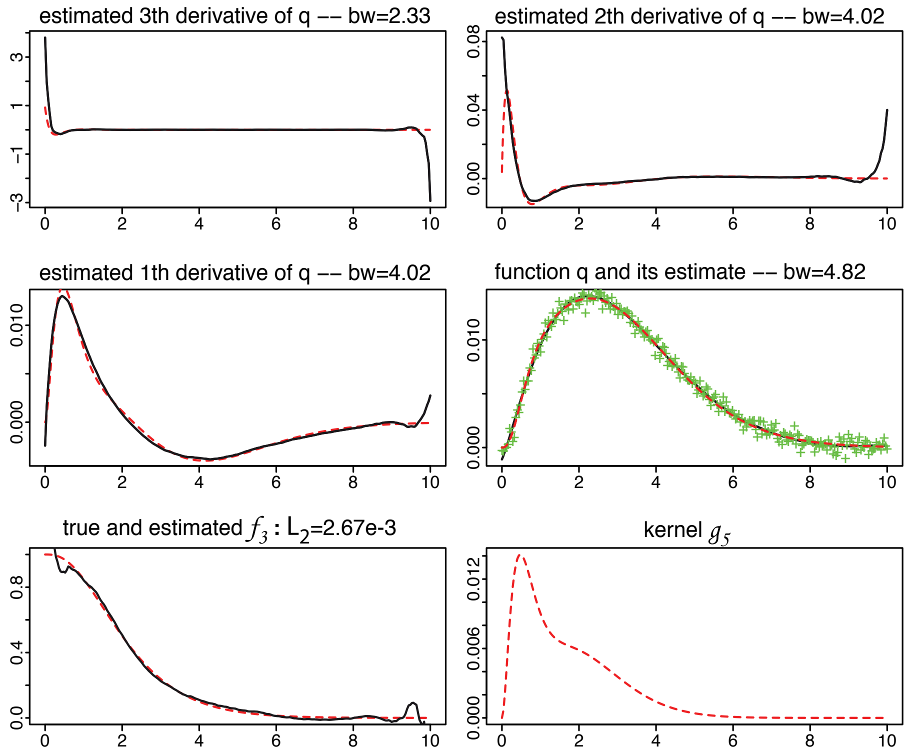

FIG 5. *Example of estimation of unknown function f_3 with kernel g_5 using $n = 250$ and $\sigma_0(g_5)/4$. Here $r = 3$. The top four sub-figures show true q and its three first derivatives $q^{(s)}$, $s = 1, 2, 3$ (dotted red lines) and their estimators (plain black lines). Selected bandwidths are specified for each estimator. Bottom left: the true function f (dotted red line) and its estimate (plain black line). Bottom right: the kernel g_5.*

4. Discussion

In the present paper, we consider Laplace deconvolution problem with discrete noisy data observed on the interval whose length T_n may increase with the sample size n. Although this problem arises in a variety of applications, to the best of our knowledge, it has been given very little attention by the statistical community. Our objective was to fill this gap and to provide statistical analysis of Laplace deconvolution problem with noisy discrete data.

The main contribution of the paper is explicit construction of a rate-optimal (in the minimax sense) Laplace deconvolution estimator which is adaptive to the regularity of the unknown function. We show that the original Laplace deconvolution problem can be reduced to nonparametric estimation of a regression function and its derivatives on the interval of growing length T_n. Although the latter problem has been well studied on a finite interval, the asymptotic increase of its length as the sample size grows raises a new challenge. Whereas the forms

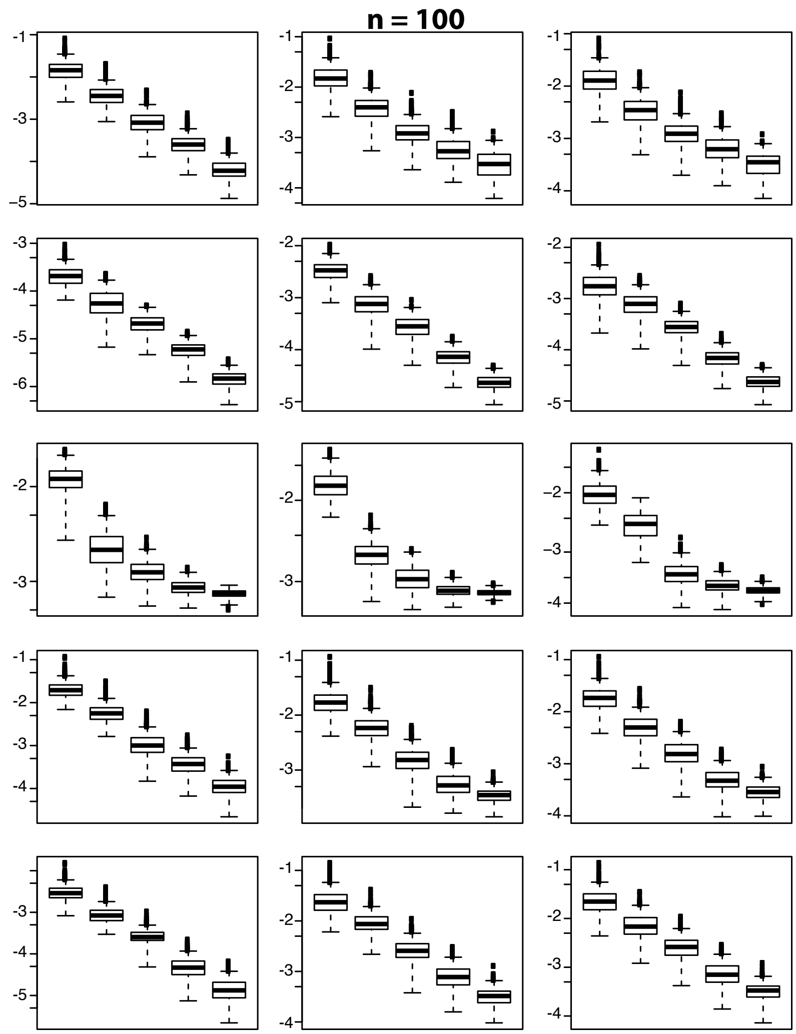

Fig 6. *Box-plots of the mean square errors for $n = 100$, 400 simulation runs and each of the triplets $(g, f, \sigma_0(g))$. Each line represents a kernel from g_1 (top) to g_5 (bottom). Each column represents an unknown function f from f_1 (left) to f_3 (right). In every sub-figure, going from left to right, each boxplot corresponds to a different noise level $\sigma_0(g_j)/2^i$ for $i = 0, \ldots, 4$. The empirical risks are presented on log-scale with the basis 10.*

of the estimators remain standard, the choices of the parameters and the minimax convergence rates, which are expressed in terms of T_n^2/n in this case, are affected by the asymptotic growth of the length of the interval.

In the present paper, we use kernel estimators with a global bandwidth adaptively chosen by the Lepski procedure (e.g., [31]) and establish asymptotic minimaxity of the resulting Laplace deconvolution estimator over a wide range of Sobolev classes. One can, however, apply other types of estimators (e.g., local

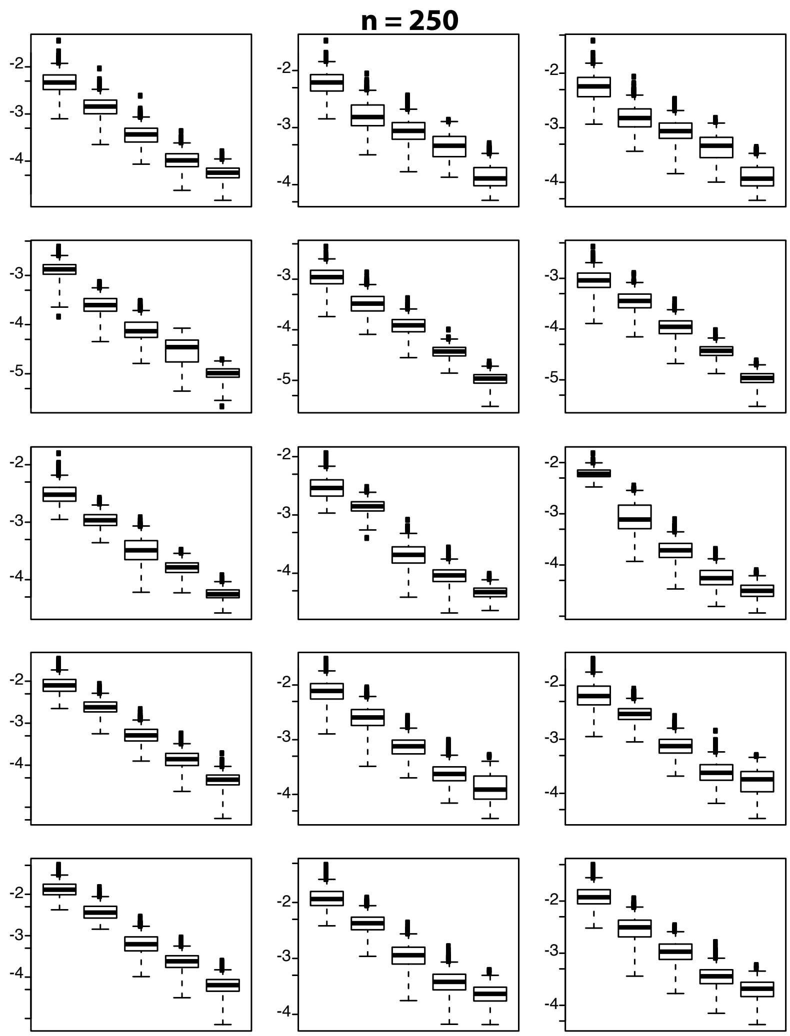

Fig 7. *Box-plots of the mean square errors for* $n = 250$. *See Figure 6 for other details.*

polynomial regression, splines or wavelets). In particular, we believe that the use of wavelet-based methods can extend the adaptive minimaxity range from Sobolev to more general Besov classes.

We illustrate the theory by examples of construction of explicit expressions for estimators of f based on observations governed by equation (1.3) with various kernels. Simulation study shows, that, in addition to providing asymptotic optimality, the proposed Laplace deconvolution estimator demonstrates good finite sample performance.

The present paper provides the first comprehensive statistical treatment of Laplace deconvolution problem, though a number of open questions remain be-

Table 1

Average (over 400 simulation runs) mean square errors and standard deviations (in parentheses) for kernels g_1 to g_5 and unknown function f_1 to f_3, for $n = 100$ (upper part) and $n = 250$ (lower part) and for the noise level equal to $\sigma_0(g_j)/2^i$, $j = 1, \dots, 5;\ i = 0, \dots, 4$.

	$\hat{R}(\hat{f})$	i=0	i=1	i=2	i=3	i=4	
g_1	f_1	1.6e-2 (9.5e-3)	4.1e-3 (2.5e-3)	9.9e-4 (6.6e-4)	2.9e-4 (1.7e-4)	7.6e-5 (4.7e-5)	$n = 100$
	f_2	1.7e-2 (1.0e-2)	4.5e-3 (2.8e-3)	1.4e-3 (8.4e-4)	6.7e-4 (4.2e-4)	3.4e-4 (1.9e-4)	
	f_3	1.5e-2 (9.6e-3)	4.0e-3 (2.6e-3)	1.4e-3 (9.0e-4)	7.4e-4 (4.4e-4)	3.6e-4 (1.7e-4)	
g_2	f_1	2.3e-3 (1.1e-3)	6.9e-4 (4.6e-4)	2.2e-4 (8.8e-5)	6.1e-5 (2.2e-5)	1.5e-5 (5.5e-6)	
	f_2	3.5e-3 (1.5e-3)	8.2e-4 (4.0e-4)	2.9e-4 (1.3e-4)	7.5e-5 (2.8e-5)	2.4e-5 (7.7e-6)	
	f_3	2.1e-3 (1.4e-3)	8.6e-4 (4.4e-4)	2.9e-4 (1.1e-4)	7.2e-5 (2.8e-5)	2.5e-5 (7.8e-6)	
g_3	f_1	1.2e-2 (3.7e-3)	2.4e-3 (1.0e-3)	1.3e-3 (3.6e-4)	8.7e-4 (1.4e-4)	7.4e-4 (7.2e-5)	
	f_2	1.4e-2 (3.8e-3)	3.5e-3 (9.2e-4)	2.2e-3 (5.0e-4)	1.7e-3 (1.9e-4)	1.6e-3 (9.1e-5)	
	f_3	1.0e-2 (3.6e-3)	4.3e-3 (1.4e-3)	1.2e-3 (3.7e-4)	8.2e-4 (1.5e-4)	7.1e-4 (7.3e-5)	
g_4	f_1	2.2e-2 (1.2e-2)	6.3e-3 (3.3e-3)	1.2e-3 (8.3e-4)	4.2e-4 (2.4e-4)	1.2e-4 (6.5e-5)	
	f_2	2.0e-2 (1.2e-2)	6.6e-3 (3.6e-3)	1.8e-3 (1.0e-3)	6.0e-4 (3.1e-4)	3.6e-4 (1.1e-4)	
	f_3	2.1e-2 (1.2e-2)	5.7e-3 (3.5e-3)	1.8e-3 (1.0e-3)	5.5e-4 (3.0e-4)	3.0e-4 (1.1e-4)	
g_5	f_1	3.2e-2 (1.7e-2)	9.4e-3 (4.6e-3)	2.8e-3 (1.2e-3)	5.3e-4 (3.2e-4)	1.6e-4 (9.9e-5)	
	f_2	2.7e-2 (1.7e-2)	9.8e-3 (4.5e-3)	2.9e-3 (1.5e-3)	8.8e-4 (4.7e-4)	3.4e-4 (1.4e-4)	
	f_3	2.6e-2 (1.7e-2)	8.1e-3 (4.8e-3)	2.9e-3 (1.5e-3)	8.3e-4 (4.8e-4)	3.4e-4 (1.4e-4)	
g_1	f_1	5.5e-3 (3.4e-3)	1.6e-3 (9.0e-4)	4.1e-4 (2.3e-4)	1.2e-4 (5.7e-5)	5.9e-5 (2.2e-5)	$n = 250$
	f_2	7.0e-3 (3.8e-3)	1.9e-3 (1.2e-3)	1.0e-3 (5.6e-4)	5.3e-4 (2.6e-4)	1.6e-4 (9.1e-5)	
	f_3	6.5e-3 (3.9e-3)	1.8e-3 (1.1e-3)	1.0e-3 (5.4e-4)	5.0e-4 (2.6e-4)	1.5e-4 (8.4e-5)	
g_2	f_1	1.4e-3 (5.2e-4)	2.7e-4 (1.2e-4)	9.1e-5 (5.4e-5)	3.5e-5 (1.9e-5)	1.0e-5 (3.1e-6)	
	f_2	1.2e-3 (5.4e-4)	3.7e-4 (1.8e-4)	1.3e-4 (5.2e-5)	3.8e-5 (1.1e-5)	1.1e-5 (3.2e-6)	
	f_3	1.0e-3 (4.9e-4)	3.9e-4 (1.8e-4)	1.2e-4 (4.8e-5)	3.8e-5 (1.1e-5)	1.1e-5 (3.4e-6)	
g_3	f_1	3.4e-3 (1.7e-3)	1.2e-3 (4.1e-4)	3.8e-4 (2.2e-4)	1.7e-4 (4.7e-5)	5.8e-5 (1.4e-5)	
	f_2	3.6e-3 (2.2e-3)	1.4e-3 (3.8e-4)	2.3e-4 (1.1e-4)	9.6e-5 (3.4e-5)	4.8e-5 (1.2e-5)	
	f_3	6.3e-3 (1.4e-3)	1.0e-3 (7.1e-4)	2.1e-4 (1.1e-4)	6.1e-5 (2.9e-5)	3.3e-5 (1.2e-5)	
g_4	f_1	8.9e-3 (4.5e-3)	2.7e-3 (1.1e-3)	5.8e-4 (3.0e-4)	1.6e-4 (8.6e-5)	4.9e-5 (2.3e-5)	
	f_2	8.8e-3 (4.6e-3)	2.9e-3 (1.5e-3)	8.2e-4 (3.8e-4)	2.8e-4 (1.5e-4)	1.6e-4 (9.6e-5)	
	f_3	7.7e-3 (4.8e-3)	3.2e-3 (1.3e-3)	8.1e-4 (3.7e-4)	2.9e-4 (1.9e-4)	1.9e-4 (9.6e-5)	
g_5	f_1	1.5e-2 (7.0e-3)	4.1e-3 (2.0e-3)	7.4e-4 (4.4e-4)	2.7e-4 (1.4e-4)	7.2e-5 (4.0e-5)	
	f_2	1.3e-2 (6.7e-3)	4.6e-3 (1.9e-3)	1.3e-3 (6.4e-4)	4.3e-4 (2.3e-4)	2.4e-4 (9.9e-5)	
	f_3	1.4e-2 (6.9e-3)	3.5e-3 (1.9e-3)	1.2e-3 (5.9e-4)	4.1e-4 (2.2e-4)	2.2e-4 (9.4e-5)	

yond its scope. In particular, an interesting challenge would be to study Laplace deconvolution with an unstable resolvent, where Assumption (A2) does not hold. Another important problem would be to study the equation (1.3) when the kernel g is not completely known and is estimated from observations.

5. Appendix

Throughout the proofs we use C to denote a generic positive constant, not necessarily the same each time it is used, even within a single equation.

Proof of Theorem 1

Although the rates are derived by standard methods described in, e.g., [43], the challenging part of the proof is constructing the set of test functions and, subsequently, producing upper bounds for the Kullback-Leibler divergence.

The main idea of the proof is to find a subset of functions $\mathcal{F} \subset W_A^m$ such that for any pair $f_1,\ f_2 \in \mathcal{F}$,

$$\|f_1 - f_2\|^2_{L_2([0,T_n])} \geq 4C(T_n^2 n^{-1})^{2m/(2(m+r)+1)} \tag{5.1}$$

and the Kullback-Leibler divergence

$$\mathbb{K}(\mathbb{P}_{f_1}, \mathbb{P}_{f_2}) = \frac{||\boldsymbol{q}_1 - \boldsymbol{q}_2||^2_{\mathbb{R}^n}}{2\sigma^2} \leq \frac{\log \mathrm{card}(\mathcal{F})}{16}, \tag{5.2}$$

where log stands for natural logarithm and vectors $\boldsymbol{q}_j$, $j = 1, 2$, have components $\boldsymbol{q}_{ji} = (g * f_j)(t_i)$, $i = 1, \ldots, n$. The result will then follow immediately from Lemma A.1 of [4]:

Lemma 1 (Bunea, Tsybakov, Wegkamp (2007), Lemma A.1)**.** *Let $\mathcal{F}$ be a set of functions of cardinality card$(\mathcal{F}) \geq 2$ such that*
(i) $\|f_1 - f_2\|^2 \geq 4\delta^2$ for any $f_1, f_2 \in \mathcal{F}$, $f_1 \neq f_2$,
(ii) the Kullback divergences $\mathbb{K}(\mathbb{P}_{f_1}, \mathbb{P}_{f_2})$ between the measures $\mathbb{P}_{f_1}$ and $\mathbb{P}_{f_2}$ satisfy the inequality $\mathbb{K}(\mathbb{P}_{f_1}, \mathbb{P}_{f_2}) \leq (1/16)\log(card(\mathcal{F}))$ for any $f_1, f_2 \in \mathcal{F}$.
Then, for some absolute positive constant C,

$$\inf_{\tilde{f}_n} \sup_{f \in \mathcal{F}} \mathbb{E}_f \|\tilde{f}_n - f\|^2 \geq C\delta^2,$$

where the infimum is taken over all estimates $\tilde{f}_n$ of f.

Without loss of generality, let us assume that the points are equally spaced, i.e. $t_i - t_{i-1} = T_n/n$, $i = 1, \ldots, n$. To construct such a subset $\mathcal{F}$, define integers $M_n \geq 8$ and $N = \left[\frac{n}{M_n}\right]$, the largest integer which does not exceed n/M_n. Let $\lambda_n = NT_n/n$ and define points $z_l = l\,\lambda_n$, $l = 0, 1, \ldots, M_n$. Note that the latter implies that points of observation $t_j = j\,T_n/n$ in equation (1.3) are related to z_l as $z_l = t_j$ where $j = Nl$ for $l = 1, \ldots, M_n$ and $j \leq NM_n$. Note also that $\frac{T_n}{2M_n} \leq \lambda_n \leq \frac{T_n}{M_n}$.

Let $k(\cdot)$ be an infinitely differentiable function with *supp*$(k) = [0, 1]$ and such that

$$\int_0^1 x^j k(x)dx = 0, j = 0, \ldots, r-1, \quad \int_0^1 x^r k(x)dx \neq 0. \tag{5.3}$$

Introduce functions

$$\varphi_j(x) = L\,\frac{\lambda_n^m}{\sqrt{T_n}}\, k\left(\frac{x - z_{j-1}}{\lambda_n}\right) \quad l = 1, \ldots, M_n,$$

where the constant $L > 0$ will be defined later. Note that φ_j have non-overlapping supports, where *supp*$(\varphi_j) = [z_{j-1}, z_j]$.

Consider the set of all binary sequences of the length $M_n \geq 8$:

$$\Omega = \left\{\boldsymbol{\omega} = (\omega_1, \ldots, \omega_{M_n}),\ \ \omega_j = \{0, 1\}\right\} = \{0, 1\}^{M_n}$$

and the corresponding subset of functions

$$\mathcal{F} = \{f_\omega : f_\omega(t) = \sum_{j=1}^{M_n} w_j \varphi_j(t),\ \omega \in \tilde{\Omega}\}. \tag{5.4}$$

Here $\tilde{\Omega} \subset \Omega$ is such that $\log_2 \mathrm{card}(\tilde{\Omega}) \geq M_n/8$ and the Hamming distance $\rho(\boldsymbol{\omega}_1, \boldsymbol{\omega}_2) = \sum_{j=1}^{M_n} \mathbb{I}\{\boldsymbol{\omega}_{1j} \neq \boldsymbol{\omega}_{2j}\} \geq M_n/8$ for any pair $\boldsymbol{\omega}_1, \boldsymbol{\omega}_2 \in \tilde{\Omega}$ (see, e.g., Lemma 2.9 of [43] for construction of $\tilde{\Omega}$).

We now need to show that $\mathcal{F}$ in (5.4) is exactly the required set. Note first that since the supports of φ_j are non-overlapping, for any $f_\omega \in \mathcal{F}$ a straightforward calculus yields

$$||f_\omega||^2_{L_2([0,T_n])} \leq \sum_{j=1}^{M_n} ||\varphi_j||^2 = L^2 \frac{\lambda_n^{2s+1}}{T_n} M_n ||k||^2 = L^2 \lambda^{2m} ||k||^2 \leq L^2 ||k||^2$$

Similarly,

$$||f_\omega^{(m)}||^2_{L_2([0,T_n])} \leq \sum_{j=1}^{M_n} ||\varphi_j^{(m)}||^2 = \frac{L^2}{T_n} m \lambda_n ||k^{(s)}||^2 = L^2 ||k^{(m)}||^2 < \infty$$

and therefore $f_\omega \in W_A^m$, where $A = L||k||_{W^m}$. Furthermore,

$$||f_{\omega_1} - f_{\omega_2}||^2_{L_2([0,T_n])} = L^2 \frac{\lambda_n^{2m+1}}{T_n} ||k||^2 \rho(\boldsymbol{\omega}_1, \boldsymbol{\omega}_2) \geq L^2 \frac{\lambda_n^{2m+1}}{T_n} \frac{M_n}{8} \geq 4C\lambda_n^{2m}$$

and (5.1) holds provided $\lambda_n \geq C(T_n^2 n^{-1})^{-1/(2(m+r)+1)}$ for some positive constant C.

To verify (5.2), note that

$$\mathbb{K}(P_1, P_2) = \frac{1}{2\sigma^2} \sum_{i=1}^{n} [q_1(t_i) - q_2(t_i)]^2 \leq \frac{1}{\sigma^2} \sum_{j=1}^{2} Q(f_j) \tag{5.5}$$

where, suppressing index j, we write

$$\begin{aligned} Q(f) &= \sum_{i=1}^{n} \left[\int_0^{t_i} g(t_i - x) f(x) dx \right]^2 \\ &= \frac{L^2 \lambda_n^{2m}}{T_n} \sum_{i=1}^{n} \left[\sum_{l=1}^{M_n} \omega_l^{(j)} \int_0^{t_i} g(t_i - x) k\left(\frac{x - z_{l-1}}{\lambda_n}\right) dx \right]^2. \end{aligned}$$

In order to obtain an upper bound for $Q(f)$ we need the following supplementary lemma, the proof of which is presented at the end of the section.

Lemma 2. *Introduce functions $K_j(x)$ using the following recursive relation*

$$K_1(x) = \int_0^x k(t)dt, \quad K_j(x) = \int_0^x K_{j-1}(t)dt, \quad j = 2, \dots, r. \tag{5.6}$$

Then, under condition (5.3), functions $K_j(x)$, $j = 1, \dots, r$, are uniformly bounded and $K_j(1) = 0$, $j = 1, \dots, r$. Moreover,

$$\int_0^{t_i} g(t_i - x) k\left(\frac{x - z_{l-1}}{\lambda_n}\right) dx = \lambda_n^r \left[B_r K_r \left(\frac{t_i - z_{l-1}}{\lambda_n}\right) \mathbb{I}(z_{l-1} \le y_i \le z_l) + \int_{\min(z_{l-1}, t_i)}^{\min(z_l, t_i)} g^{(r)}(t_i - x) K_r \left(\frac{x - z_{l-1}}{\lambda_n}\right) dx \right]. \tag{5.7}$$

Applying equation (5.7) to the integral in $Q(f)$, obtain

$$Q(f) \le 2L^2 \lambda_n^{2m+2r} T_n^{-1} (\Delta_1 + \Delta_2) \tag{5.8}$$

where

$$\begin{aligned} \Delta_1 &= \sum_{i=1}^n \left[\sum_{l=1}^{M_n} B_r \; K_r \left(\frac{t_i - z_{l-1}}{\lambda_n}\right) \mathbb{I}(z_{l-1} \le y_i \le z_l) \right]^2, \\ \Delta_2 &= \sum_{i=1}^n \left[\sum_{l=1}^{M_n} \int_{\min(z_{l-1}, t_i)}^{\min(z_l, t_i)} g^{(r)}(t_i - x) K_r \left(\frac{x - z_{l-1}}{\lambda_n}\right) dx \right]^2. \end{aligned}$$

Observe that for any t and any l_1 and l_2 such that $l_1 \neq l_2$, one has $K_r(\lambda_n^{-1}(t - z_{l_1})) K_r(\lambda_n^{-1}(t - z_{l_2})) = 0$. Also, for each i, $K_r(\lambda_n^{-1}(ti - z_l)) \neq 0$ for only one value of l, namely, for $l = [i/N] + 1$ where $[x]$ is the largest integer which does not exceed x. Therefore,

$$\Delta_1 \le B_r^2 \sum_{i=1}^n K_r^2 \left(\frac{t_i - z_{[i/N]}}{\lambda_n}\right) \le n \; B_r^2 \|K_r\|_\infty^2, \tag{5.9}$$

where $\|\cdot\|_\infty$ is the supremum norm. In order to obtain an upper bound for Δ_2, observe that for any nonnegative function $F(x)$ one has

$$\int_{\min(z_{l-1}, t_i)}^{\min(z_l, t_i)} F(x) dx \le \int_{z_{l-1}}^{z_l} F(x) dx.$$

Hence, we derive

$$\begin{aligned} \Delta_2 &\le \sum_{i=1}^n \left[\sum_{l=1}^{M_n} \int_{z_{l-1}}^{z_l} \left| g^{(r)}(t_i - x) K_r \left(\frac{x - z_{l-1}}{\lambda_n}\right) \right| dx \right]^2 \\ &\le \sum_{i=1}^n \|K_r\|_\infty^2 \left[\sum_{l=1}^{M_n} \int_{z_{l-1}}^{z_l} |g^{(r)}(t_i - x)| dx \right]^2 \le n \, \|g^{(r)}\|^2 \|K_r\|_\infty^2. \end{aligned} \tag{5.10}$$

Combining formulae (5.5)–(5.10), we obtain that, in order to satisfy the condition (5.2), we need the following inequality to hold

$$\mathbb{K}(\mathbb{P}_{f_1}, \mathbb{P}_{f_2}) \le \frac{2L^2 \lambda_n^{2m+2r} n}{\sigma^2 T_n} \|K_r\|_\infty^2 [B_r^2 + \|g^{(r)}\|_2^2] \le \frac{1}{16} \frac{M_n \log 2}{8}. \tag{5.11}$$

Note that $\frac{T_n}{M_n}\left(1 - \frac{T_n}{M_n}\right) \le \lambda_n \le \frac{T_n}{M_n}$. Choosing $M_n = Cn^{1/(2(m+r)+1)} \times T_n^{(2(m+r)-1)/(2(m+r)+1)}$ and observing that $T_n/M_n \to 0$ as $n \to \infty$, obtain $\lambda_n \ge T_n/(2M_n) \ge C(T_n^2 n^{-1})^{1/(2(m+r)+1)}$. Therefore, both conditions (5.1) and (5.2) hold and theorem is proved. □

Proof of Theorem 2

To prove Theorem 2 we use the following Lemma 3 which can be viewed as a version of Theorem 7.2.4 of [25] adapted to our notations.

Lemma 3. *Let s_g be such that*

$$\inf_{Re(s)=s_g} |\tilde{g}(s)| > 0 \quad \textit{and} \quad \lim_{\substack{|s|\to\infty \\ Re(s)\ge s_g}} |s^r \tilde{g}(s)| > 0. \tag{5.12}$$

Then, solution $\phi(\cdot)$ of equation (2.4) can be presented as

$$\phi(t) = \sum_{l=0}^{L} \sum_{j=0}^{\alpha_l - 1} \frac{a_{l,j}}{j!} t^j e^{s_l t} + \phi_1(t) \tag{5.13}$$

where L is the total number of distinct zeros s_l of $s^r\tilde{g}(s)$ such that $Re(s_l) > Re(s_g)$, α_l is the order of zero s_l and $\phi_1 \in L_1$.

Choose s_g such that $s^* < s_g < 0$. Then, the first condition in (5.12) immediately follows from Assumption (A2). To validate the second assumption in (5.12), note that for $s = s_1 + is_2$ conditions $Re(s) \ge s_g$ and $|s| \to \infty$ imply that either $s_1 \to \infty$ or $|s_2| \to \infty$, or both. Recall that $s^r\tilde{g}(s) = B_r + \widetilde{G^{(r)}}(s)$. If $s_1 \to \infty$, no matter whether s_2 is finite or $s_2 \to \infty$, one has

$$\lim_{Re(s)\to\infty} |s^r \tilde{g}(s)| = \lim_{Re(s)\to\infty} |B_r + \int_0^\infty g^{(r)}(t) e^{-st} dt| = |B_r| > 0. \tag{5.14}$$

If s_1 is finite, $s_1 \ge s_g$, and $|s_2| \to \infty$, then Laplace transform $\widetilde{g^{(r)}}(s) = \int_0^\infty g^{(r)}(t)e^{-st}dt$ is equal to Fourier transform $\mathcal{F}[g^{(r)}(t)e^{-s_1 t}](s_2)$ of function $g^{(r)}(t)e^{-s_1 t}$ at the point s_2. Since $g^{(r)}(t)e^{-s_1 t} \in L_1(\mathbb{R}^+)$, one obtains

$$\lim_{|s_2|\to\infty} \int_0^\infty g^{(r)}(t)e^{-st}dt = \lim_{|s_2|\to\infty} \mathcal{F}[g^{(r)}(t)e^{-s_1 t}](s_2) = 0,$$

and (5.14) holds again. Hence, the second assumption in (5.12) is valid, and Lemma 3 can be applied.

Note that, under Assumption (A2), $\tilde{g}(s)$ has no zeros with $Re(s) > s_g$ and, therefore, $s^r\tilde{g}(s)$ has a single zero of r-th order at $s = 0$. Lemma 3 yields then that $\phi(t) = \phi_0(t) + \phi_1(t)$, where

$$\phi_0(t) = \sum_{j=0}^{r-1} \frac{a_{0,j}}{j!} t^j, \quad a_{0,j} = \phi^{(j)}(0), \tag{5.15}$$

and integrating by parts, one has

$$\int_0^t q^{(r)}(t-\tau)\phi_0(\tau)d\tau = \sum_{j=0}^{r-1} \phi_0^{(r-j-1)}(0)q^{(j)}(t), \tag{5.16}$$

that completes the proof of (2.8).

In order to prove (2.9) – (2.12), note that it follows from equation (2.6) that $\tilde{\phi}(s)$ has poles s_l, $l = 0, \ldots, M$, of respective orders α_l, where $s_0 = 0$ and $\alpha_0 = r$. Since, by (5.14), one has

$$\lim_{\substack{|s|\to\infty \\ Re(s)\geq s_g}} |s^r \tilde{g}(s)| > 0$$

and, therefore, $\tilde{\phi}$ does not have a pole at infinity. Then, $\tilde{\phi}$ is a rational function and, consequently, can be represented using Cauchy integral formula

$$\tilde{\phi}(s) = -\frac{1}{2\pi i}\sum_{l=0}^{M} \oint_{C_l} \frac{\tilde{\phi}(z)}{z-s} dz$$

where C_l, $l = 0, \ldots, M$, is a circle around the pole s_l such that this circle does not enclose any other pole of $\tilde{\phi}$ (see [30], Section 5.14). Using Laurent expansion of $\tilde{\phi}(z)$ around s_l, we have

$$\begin{aligned} I_l(s) &= \frac{1}{2\pi i}\oint_{C_l} \frac{\tilde{\phi}(z)}{z-s} dz \\ &= -\sum_{j=0}^{\alpha_l -1} \frac{1}{(s-s_l)^{j+1}} \; \frac{1}{(\alpha_l - 1 - j)!} \; \frac{d^{\alpha_l - j-1}}{ds^{\alpha_l - j - 1}} \left[(s-s_l)^{\alpha_l} \tilde{\phi}(s) \right] \Bigg|_{s=s_l} \end{aligned}$$

Combining the last two expressions and taking inverse Laplace transform of $\tilde{\phi}(s)$ yields

$$\phi(t) = \sum_{l=0}^{M} \sum_{j=0}^{\alpha_l -1} \frac{a_{l,j}}{j!} t^j e^{s_l t} = \phi_0(t) + \phi_1(t),$$

where ϕ_0 is given by (5.15), same as before, and

$$\phi_1(t) = \sum_{l=1}^{M} \sum_{j=0}^{\alpha_l -1} \frac{a_{l,j}}{j!} t^j e^{s_l t}. \tag{5.17}$$

Repeat calculations in (5.16) and also note that, by similar considerations, for every $j = 0, \ldots, \alpha_l - 1$, one can write

$$\begin{aligned} \int_0^t q^{(r)}(t-x) x^j e^{s_l x} dx &= \sum_{k=0}^{r-1} q^{(r-k-1)}(t) \, \frac{d^k}{dx^k} \left[x^j e^{s_l x} \right] \Bigg|_{x=0} \\ &\quad + \int_0^t q(t-x) \frac{d^{(r-1)}}{dx^{(r-1)}} \left[x^j e^{s_l x} \right] dx. \end{aligned}$$

To complete the proof, evaluate the derivatives, observe that

$$\frac{d^k}{dx^k}\left[x^j e^{s_l x}\right]\Bigg|_{x=0} = \binom{k}{j} s_l^{k-j}$$

and interchange summation with respect to j and k. □

Proof of Theorem 3

Since the estimator (2.14) is just a particular form of the estimator (2.13), it is sufficient to carry out the proof for the estimator (2.13) of f. From (2.8), one immediately obtains

$$E||\hat{f}_n - f||^2_{[0,T_n]} \le \frac{r+2}{B_r^2}\Bigg(E||\widehat{q^{(r)}_{\hat{\lambda}_{r,n}}} - q^{(r)}||^2_{[0,T_n]} + \sum_{j=0}^{r-1} a^2_{0,r-1-j} E||\widehat{q^{(j)}_{\hat{\lambda}_{j,n}}} - q^{(j)}||^2_{[0,T_n]} + ||\widehat{q^{(r)}_{\hat{\lambda}_{r,n}}} * \phi_1 - q^{(r)} * \phi_1||^2_{[0,T_n]}\Bigg), \tag{5.18}$$

where $\widehat{q^{(j)}_{\hat{\lambda}_{j,n}}}$, $j = 0, \dots, r$ are given in (2.19).

The proof is based on the following proposition which provides upper bounds for the risks $E||\widehat{q^{(j)}_{\hat{\lambda}_{j,n}}} - q^{(j)}||^2_{[0,T_n]}$, $j = 0, \dots, r$ in (5.18).

Proposition 1. *Let condition (2.1) and Assumptions (A1)-(A4) hold. Let kernel K_j be of order (L, j), where $L > r$ and $0 \le j \le r$, and satisfies Assumptions (K1) and (K2). Then, for all $A' > 0$,*

$$\sup_{q \in W^{m+r}_{A'}} E||\widehat{q^{(j)}_{\hat{\lambda}_{j,n}}} - q^{(j)}||^2_{[0,T_n]} = O\left(\left(\frac{T_n^2}{n}\right)^{\frac{2(r+m-j)}{2(r+m)+1}}\right). \tag{5.19}$$

In particular, Proposition 1 implies that the errors of estimating $q^{(j)}$ in (5.18) are dominated by the estimation error of the highest order derivative $q^{(r)}$. Furthermore, $\phi_1 \in L_1$ (see Theorem 2) and, therefore,

$$\|\widehat{q^{(r)}_{\hat{\lambda}_{r,n}}} * \phi_1 - q^{(r)} * \phi_1\|_2 \le \|\phi_1\|_1 \cdot \|\widehat{q^{(r)}_n} - q^{(r)}\|_2 = O\left(\|\widehat{q^{(r)}_{\hat{\lambda}_{r,n}}} - q^{(r)}\|_2\right)$$

(see also Theorem 2.2.2 of [25]).

Thus, (5.18) and Proposition 1 yield

$$E||\hat{f}_n - f||^2_{[0,T_n]} = O\left(E||\widehat{q^{(r)}_{\hat{\lambda}_{r,n}}} - q^{(r)}||^2_{[0,T_n]}\right) = O\left(\left(\frac{T_n^2}{n}\right)^{\frac{2m}{2m+2r+1}}\right)$$

□

Proof of Proposition 1

For simplicity of notations we drop the index n in $\hat{\lambda}_{j,n}$.

Recall that under Assumptions (A1)-(A3), $q \in W^{r+m}$ (see Section 2.3). By the standard asymptotic calculus for kernel estimation (see, e.g., [18]) for estimator (2.15) and any interior point t of $(0, T_n)$, one then has

$$\begin{aligned} Var\left(\widehat{q_{\lambda}^{(j)}}(t)\right) &= \frac{\sigma^2}{\lambda_j^{2(j+1)}} \sum_{i=1}^{n} (t_i - t_{i-1})^2 K_j^2\left(\frac{t_i - t}{\lambda_j}\right) \\ &= \frac{\sigma^2}{\lambda_j^{2j+1}} \frac{T_n}{n} \int K_j^2(u) du \, (1 + o(1)). \end{aligned}$$

The required boundary corrections ensure the same order of error for the values t close to the boundaries ([18]) and the integrated variance then is

$$V_j(\lambda_j) = \int_0^{T_n} Var\left(\widehat{q_{\lambda_j}^{(j)}}(t)\right) dt = V_{0j} \, \frac{T_n^2}{\lambda_j^{2j+1} n} \, (1 + o(1)), \tag{5.20}$$

where $V_{0j} = \sigma^2 ||K_j||^2$. Similarly, the integrated squared bias can be written as

$$B_j^2(\lambda_j, q) = \int_0^{T_n} \left(E\left(\widehat{q_{\lambda_j}^{(j)}}(t)\right) - q^{(j)}(t)\right)^2 dt = B_{0j} \lambda_j^{2(r+m-j)} (1 + o(1)), \tag{5.21}$$

where $B_{0j} = B_0^{-1} ||q^{(r+m)}||^2 ||K_j||^2$ and $B_0 = 2\left((r+m-1)!\right)^2 (2(r+m) - 1)(2(r+m)+1)$. Hence, for any radius $A' > 0$,

$$\begin{aligned} \sup_{q \in W_{A'}^{m+r}} E||\widehat{q_{\lambda_j}^{(j)}} - q^{(j)}||^2_{L_2([0,T_n)]} &= \sup_{q \in W_{A'}^{m+r}} \left(V_j(\lambda_j) + B_j^2(\lambda_j, q)\right) \\ &= O\left(\frac{T_n^2}{\lambda_j^{2j+1} n}\right) + O\left(\lambda_j^{2(r+m-j)}\right). \end{aligned} \tag{5.22}$$

It follows from (5.20) and (5.21) that the asymptotically optimal bandwidth that minimizes $E||\widehat{q_{\lambda_j}^{(j)}} - q^{(j)}||^2_{L_2([0,T_n])}$ is

$$\lambda_j^* = O\left(\left(\frac{T_n^2}{n}\right)^{\frac{1}{2(r+m)+1}}\right) \tag{5.23}$$

and the corresponding risk of estimating $q^{(j)}$ is given by

$$\sup_{q \in W_{A'}^{m+r}} E||\widehat{q_{\lambda_j^*}^{(j)}} - q^{(j)}||^2_{L_2([0,T_n])} = O\left(\left(\frac{T_n^2}{n}\right)^{\frac{2(r+m-j)}{2(r+m)+1}}\right). \tag{5.24}$$

Now we need to prove that (5.24) remains valid when λ_j^* is replaced by $\hat{\lambda}_j$ selected by Lepski procedure, that is,

$$\sup_{q\in W_{A'}^{m+r}} E||\widehat{q_{\hat{\lambda}_j}^{(j)}} - q^{(j)}||^2 = O\left(\left(\frac{T_n^2}{n}\right)^{\frac{2(r+m-j)}{2(r+m)+1}}\right)$$

for all $A' > 0$. Set d_j and λ_j^* in (5.23) to be, respectively,

$$d_j = \frac{C_j - \mu||K_j||}{2||K_j||}, \quad \lambda_j^* = \left(d_j^2 \frac{\sigma^2 B_0}{2(A')^2} \frac{T_n^2}{n}\right)^{\frac{1}{2(r+m)+1}},$$

where C_j is defined in (2.18). Note that

$$\begin{aligned} E\|\widehat{q_{\hat{\lambda}_j}^{(j)}} - q^{(j)}\|^2 &= E\left\{\|\widehat{q_{\hat{\lambda}_j}^{(j)}} - q^{(j)}\|^2 I(\hat{\lambda}_j \geq \lambda_j^*)\right\} + E\left\{\|\widehat{q_{\hat{\lambda}_j}^{(j)}} - q^{(j)}\|^2 I(\hat{\lambda}_j < \lambda_j^*)\right\} \\ &= \Delta_1 + \Delta_2. \end{aligned}$$

For $\hat{\lambda}_j \geq \lambda_j^*$, equations (5.24) and (2.17) imply that uniformly over $q \in W_{A'}^{m+r}$

$$\begin{aligned} \Delta_1 &\leq 2E\left\{\|\widehat{q_{\hat{\lambda}_j}^{(j)}} - \widehat{q_{\lambda_j^*}^{(j)}}\|^2 I(\hat{\lambda}_j > \lambda_j^*)\right\} + 2E\left\{\|\widehat{q_{\lambda_j^*}^{(j)}} - q^{(j)}\|^2 I(\hat{\lambda}_j > \lambda_j^*)\right\} \\ &= O\left(n^{-1}T_n^2(\lambda_j^*)^{-(2j+1)}\right) + O\left((n^{-1}T_n^2)^{-\frac{2(r+m-j)}{2(r+m)+1}}\right) \\ &= O\left((n^{-1}T_n^2)^{-\frac{2(r+m-j)}{2(r+m)+1}}\right). \end{aligned} \tag{5.25}$$

For $(n^{-1}T_n^2)^{\frac{1}{2j+1}} \leq \hat{\lambda}_j < \lambda_j^*$, by direct calculus similar to that carried out above, one can show that

$$\sup_{q\in W_{A'}^{m+r}} E\|\widehat{q_{\hat{\lambda}_j}^{(j)}} - q^{(j)}\|^4 = O\left((\lambda_j^*)^{-2(2j+1)} n^{-2} T_n^4\right) + O((\lambda_n^*)^{4(r+m-j)}) = O(1).$$

Hence,

$$\begin{aligned} \sup_{q\in W_{A'}^{m+r}} \Delta_2 &\leq \sup_{q\in W_{A'}^{m+r}} \sqrt{E\|\widehat{q_{\hat{\lambda}_j}^{(j)}} - q^{(j)}\|^4}\sqrt{P(\hat{\lambda}_j < \lambda_j^*)} \\ &= O\left(\sup_{q\in W_{A'}^{m+r}} \sqrt{P(\hat{\lambda}_j < \lambda_j^*)}\right). \end{aligned} \tag{5.26}$$

If $\lambda_j^* > \hat{\lambda}_j$, it follows from definition (2.17) of $\hat{\lambda}_j$ that there exists $\tilde{h} < \lambda_j^*$ such that $\|\widehat{q_{\lambda_j^*}^{(j)}} - \widehat{q_{\tilde{h}}^{(j)}}\|^2 > 4\,C_j^2 n^{-1}\sigma^2 T_n^2 \tilde{h}^{-(2j+1)}$, where, by (2.18) and definition of d_j, we have $C_j = \|K_j\|(\mu + 2d_j)$. It follows from (5.20) and (5.21) that, for all $h < \lambda_j^*$, the variance term dominates the squared bias, that is,

$$\sup_{q\in W_{A'}^{m+r}} \|E\widehat{q_h^{(j)}} - q^{(j)}\|^2 \leq d_j^2\sigma^2\|K_j\|^2 n^{-1} T_n^2 h^{-(2j+1)}.$$

Hence, for all $\tilde{h} < \lambda_j^*$ and $q \in W_{A'}^{m+r}$, one has

$$
\begin{aligned}
& P\left(\|\widehat{q_{\lambda_j^*}^{(j)}} - \widehat{q_{\tilde{h}}^{(j)}}\|^2 > 4\, C_j^2 n^{-1}\sigma^2 T_n^2 \tilde{h}^{-(2j+1)}\right) \\
& \quad < \quad P\left(\|\widehat{q_{\lambda_j^*}^{(j)}} - E q_{\lambda_j^*}^{(j)}\|^2 > \sigma^2 \|K_j\|^2 (\mu + d_j)^2 n^{-1} T_n^2 \tilde{h}^{-(2j+1)}\right) \\
& \qquad + P\left(\|\widehat{q_{\tilde{h}}^{(j)}} - E q_{\tilde{h}}^{(j)}\|^2 > \sigma^2 \|K_j\|^2 (\mu + d_j)^2 n^{-1} T_n^2 \tilde{h}^{-(2j+1)}\right)
\end{aligned}
$$

due to $C_j - \|K_j\| d_j > \|K_j\|(\mu + d_j)$. Thus, uniformly over $q \in W_{A'}^{s+r}$, one has

$$
P(\hat{\lambda}_j < \lambda_j^*) \leq \sum_{\substack{h \in \Lambda_j \\ h \leq \lambda_j^*}} P(\tilde{h} = h)\; P\left(\|\widehat{q_{\lambda_j^*}^{(j)}} - \widehat{q_h^{(j)}}\|^2 > 4\, \sigma^2 C_j^2 n^{-1} T_n^2 h^{-(2j+1)}\right) \quad (5.27)
$$

$$
\leq 2 \sum_{\substack{h \in \Lambda_j \\ h \leq \lambda_j^*}} P(\tilde{h} = h)\; P\left(\|\widehat{q_h^{(j)}} - \widehat{E q_h^{(j)}}\|^2 \geq \sigma^2 \|K_j\|^2 (\mu + d_j)^2 n^{-1} T_n^2 h^{-(2j+1)}\right).
$$

Note that

$$
\|\widehat{q_h^{(j)}} - \widehat{E q_h^{(j)}}\|^2 = \left\| \sum_{i=1}^n h^{-(j+1)} K_j\left(\frac{t - t_i}{h}\right)(t_i - t_{i-1})\epsilon_i \right\|^2 = h^{-(2j+1)} n^{-2} T_n^2\, \boldsymbol{\epsilon}^T \boldsymbol{Q} \boldsymbol{\epsilon},
$$

where $\boldsymbol{Q}$ is an $n \times n$ symmetric nonnegative-definite matrix with elements

$$
Q_{il} = \frac{n^2}{T_n^2}(t_i - t_{i-1})(t_l - t_{l-1}) \int_{-1}^{1} K_j(z) K_j\left(z + \frac{t_i - t_l}{h}\right) dz. \quad (5.28)
$$

Then,

$$
\begin{aligned}
& P\left(\|\widehat{q_h^{(j)}} - \widehat{E q_h^{(j)}}\|^2 \geq \sigma^2 \|K_j\|^2 (\mu + d_j)^2 n^{-1} T_n^2 h^{-(2j+1)}\right) \\
& \qquad = P\left(\boldsymbol{\epsilon}^T \boldsymbol{Q} \boldsymbol{\epsilon} \geq n\, \sigma^2 \|K_j\|^2 (\mu + d_j)^2\right). \quad (5.29)
\end{aligned}
$$

Applying a χ^2-type inequality which initially appeared in [29], was improved by [8] and furthermore by [21], we derive that, for any $x > 0$,

$$
P\left(\sigma^{-2} \boldsymbol{\epsilon}^T \boldsymbol{Q} \boldsymbol{\epsilon} \geq \left[\sqrt{\mathrm{Tr}(\boldsymbol{Q})} + \sqrt{x \rho_{\max}^2(\boldsymbol{Q})}\right]^2\right) \leq e^{-x}, \quad (5.30)
$$

where $\mathrm{Tr}(\boldsymbol{Q})$ is the trace of $\boldsymbol{Q}$, and $\rho_{\max}^2(\boldsymbol{Q})$ is the maximal eigenvalue of $\boldsymbol{Q}$. Note that

$$
\mathrm{Tr}(\boldsymbol{Q}) = \frac{n^2}{T_n^2} \sum_{i=1}^n (t_i - t_{i-1})^2 \|K_j\|^2 \leq n \mu^2 \|K_j\|^2,
$$

and $\rho^2_{\max}(\boldsymbol{Q})$ is the spectral norm of matrix $\boldsymbol{Q}$ which is dominated by any other norm. In particular,

$$\begin{aligned}\rho^2_{\max}(\boldsymbol{Q}) &\le \max_k \sum_{l=1}^n |Q_{kl}| \\ &= \frac{n^2}{T_n^2}\max_k (t_k - t_{k-1}) \int_{-1}^1 |K_j(z)| \left[\sum_{l=1}^n \left|K_j\left(z + \frac{t_k - t_l}{h}\right)\right| (t_l - t_{l-1})\right] dz.\end{aligned}$$

Since

$$\begin{aligned}\sum_{l=1}^n \left|K_j\left(z + \frac{t_k - t_l}{h}\right)\right| (t_l - t_{l-1}) &= \int_{-1}^1 \left|K_j\left(z + \frac{t_k - t}{h}\right)\right| dt(1+o(1)) \\ &= h\int_{-1}^1 \left|K_j\left(z + \frac{t_k}{h} - y\right)\right| dt(1+o(1)),\end{aligned}$$

we derive

$$\begin{aligned}\rho^2_{\max}(\boldsymbol{Q}) &\le \frac{n^2}{T_n^2}\max_k \left[(t_k - t_{k-1})\, h \int_{-1}^1\int_{-1}^1 |K_j(z)||K_j(z + t_k/h - y)|dzdy\right] \\ &\le \mu\frac{nh}{T_n}\left[\int_{-1}^1 |K_j(z)|dz\right]^2 \le 2\mu\|K_j\|^2\,\frac{nh}{T_n}.\end{aligned}$$

Using inequality (5.30) with $x = d_j^2 T_n/(2\mu h)$ and $h < \lambda_j^*$ one obtains

$$\begin{aligned}P\left(\|\widehat{q_h^{(j)}} - E\widehat{q_h^{(j)}}\|^2 \ge \frac{\sigma^2\|K_j\|^2(\mu + d_j)^2 T_n^2}{nh^{2j+1}}\right) &\le \exp\left(-\frac{d_j^2 T_n}{2\mu h}\right) \\ &\le \exp\left(-c_j n^{\frac{1}{2(r+m)+1}} T_n^{\frac{2(r+m)-1}{2(r+m)+1}}\right) \qquad (5.31)\end{aligned}$$

where c_j depends on m, A', μ and d_j. Combination of (5.25), (5.26), (5.27) and (5.31) completes the proof. □

Proof of Lemma 2. Definitions (5.6) imply that $k(x) = K_1'(x)$, $K_{j-1}'(x) = K_j(x)$ and $K_j(0) = 0$, $j = 1, \dots, r$. Observe that condition $K_j(1) = 0$, $j = 1, \dots, r$, is equivalent to

$$\int_0^1 K_j(x)dx = 0, \quad j = 0, \dots, r-1, \qquad (5.32)$$

where $K_0(x) = k(x)$. It is easy to see that (5.32) is valid for $j = 0$. For $j \ge 1$, note that, by formula (4.631) of [24],

$$K_j(x) = \int_0^x dz_{j-1}\int_0^{z_{j-1}} dz_{j-2}\dots\int_0^{z_1} k(z)dz = \frac{1}{(j-1)!}\int_0^x (x-z)^{j-1}k(z)dz. \qquad (5.33)$$

Then, for any $x \in [0,1]$, one has $|K_j(x)| \le [(j-1)!]^{-1} \|k\|_\infty \int_0^x (x-z)^{j-1} dz \le \|k\|_\infty$. Moreover, by (5.33), for $j = 1, \ldots, r-1$, one has

$$\int_0^1 K_j(x)dx = \frac{1}{(j-1)!} \int_0^1 dx \int_0^x (x-z)^{j-1} k(z)dz$$
$$= \frac{1}{(j-1)!} \int_0^1 k(z)dz \int_z^1 (x-z)^{j-1}dx = \frac{1}{(j-1)!j!} \int_0^1 (1-z)^j \, k(z)dz = 0.$$

Now, it remains to prove formula (5.7). Note that support of the function $k(u/\lambda_n - (l-1))$ coincides with (z_{l-1}, z_l), so that

$$I(i,l) = \int_0^{t_i} g(t_i - x) k\left(\frac{x - z_{l-1}}{\lambda_n}\right) dx = \int_{\min(z_{l-1},t_i)}^{\min(z_l,t_i)} g(t_i - x) k\left(\frac{x - z_{l-1}}{\lambda_n}\right) dx. \tag{5.34}$$

Formula (5.34) implies that $I(i,l) = 0$ whenever $z_{l-1} \ge y_i$. If $z_{l-1} < y_i \le z_l$, it follows from (5.34) that

$$I(i,l) = \int_{z_{l-1}}^{t_i} g(t_i - x) k\left(\frac{x - z_{l-1}}{\lambda_n}\right) dx.$$

Introduce new variable $t = x - z_{l-1}$ and denote $u_{il} = t_i - z_{l-1}$. Then, recalling condition (2.1) and using integration by parts, we derive

$$I(i,l) = \int_0^{u_{il}} g(u_{il} - t) k\left(\frac{t}{\lambda_n}\right) dt = \lambda_n g(u_{il} - t) K_1\left(\frac{t}{\lambda_n}\right)\Bigg|_0^{u_{il}}$$
$$+ \lambda_n \int_0^{u_{il}} g'(u_{il} - t) K_1\left(\frac{t}{\lambda_n}\right) dt$$
$$= \cdots = \lambda_n^r g^{(r-1)}(u_{il} - t) K_r\left(\frac{t}{\lambda_n}\right)\Bigg|_0^{u_{il}} + \lambda_n^r \int_0^{u_{il}} g^r(u_{il} - t) K_r\left(\frac{t}{\lambda_n}\right) dt.$$

Changing variables back to x, we arrive at

$$I(i,l) = \lambda_n^r \left[B_r \, K_r\left(\frac{t_i - z_{l-1}}{\lambda_n}\right) + \int_{z_{l-1}}^{t_i} g^{(r)}(t_i - x) K_r\left(\frac{x - z_{l-1}}{\lambda_n}\right) dx \right]. \tag{5.35}$$

Finally, consider the case when $z_l \le y_i$. Then, using relation $z_l = z_{l-1} + \lambda_n$, integration by parts and the fact that $K_j(0) = K_j(1) = 0$ for $j = 1, \ldots, r$, we obtain

$$I(i,l) = \int_{z_{l-1}}^{z_l} g(t_i - x) k\left(\frac{x - z_{l-1}}{\lambda_n}\right) dx = \lambda_n \int_0^1 g(t_i - z_{l-1} - \lambda_n t) k(t) dt$$
$$= \cdots = \lambda_n^{r+1} \int_0^1 g^r(t_i - z_{l-1} - \lambda_n t) K_r(t) dt$$
$$= \lambda_n^r \int_{z_{l-1}}^{z_l} g^{(r)}(t_i - x) K_r\left(\frac{x - z_{l-1}}{\lambda_n}\right) dx$$

which, in combination with (5.35), completes the proof. □

References

[1] Ameloot, M. and Hendrickx, H. (1983). Extension of the performance of Laplace deconvolution in the analysis of fluorescence decay curves. *Biophys. Journ.*, **44**, 27–38.

[2] Ameloot, M., Hendrickx, H., Herreman, W., Pottel, H., Van Cauwelaert, F. and van der Meer, W. (1984). Effect of orientational order on the decay of the fluorescence anisotropy in membrane suspensions. Experimental verification on unilamellar vesicles and lipid/alpha-lactalbumin complexes. *Biophys. Journ.*, **46**, 525–539.

[3] Bisdas, S., Konstantinou, G.N., Lee, P.S., Thng, C.H., Wagenblast, J., Baghi, M. and Koh, T.S. (2007). Dynamic contrast-enhanced CT of head and neck tumors: perfusion measurements using a distributed-parameter tracer kinetic model. Initial results and comparison with deconvolution- based analysis. *Physics in Medicine and Biology*, **52**, 6181–6196.

[4] Bunea, F., Tsybakov, A. and Wegkamp, M.H. (2007). Aggregation for Gaussian regression. *Ann. Statist.* **35**, 1674–1697. MR2351101

[5] Cao, M.,Liang, Y., Shen, C., Miller, K.D. and Stantz, K.M. (2010). Developing DCE-CT to quantify intra-tumor heterogeneity in breast tumors with differing angiogenic phenotype. *IEEE Trans. Medic. Imag.*, **29**, 1089–1092.

[6] Carroll, R. J. and Hall, P. (1988). Optimal rates of convergence for deconvolving a density. *J. Amer. Statist. Assoc.* **83**, 1184–1186.

[7] Cinzori, A.C. and Lamm, P.K. (2000). Future polynomial regularization of ill-posed Volterra equations. *SIAM J. Numer. Anal.*, **37**, 949–979. MR1749244

[8] Comte, F. (2001) Adaptive estimation of the spectrum of a stationary Gaussian sequence. *Bernoulli*, **7**, 267–298. MR2331262

[9] Comte, F., Rozenholc, Y. and Taupin, M.L. (2006). Penalized contrast estimator for adaptive density deconvolution. *Canad. J. Statist.*, **3**, 431–452. MR2328553

[10] Comte, F., Rozenholc, Y. and Taupin, M.L. (2007). Finite sample penalization in adaptive density deconvolution. *J. Stat. Comput. Simul.*, **77**, 977–1000. MR2416478

[11] Cuenod, C.A., Fournier, L., Balvay, D. and Guinebretire, J.M. (2006). Tumor angiogenesis: pathophysiology and implications for contrast-enhanced MRI and CT assessment. *Abdom. Imaging*, **31**, 188-193.

[12] Cuenod, C-A., Favetto, B., Genon-Catalot, V., Rozenholc, Y. and Samson, A. (2011). Parameter estimation and change-point detection from Dynamic Contrast Enhanced MRI data using stochastic differential equations. *Math. Biosci.*, **233**-1, 68–76. MR2865652

[13] Delaigle, A., Hall, P. and Meister, A. (2008). On deconvolution with repeated measurements. *Ann. Statist.*, **36**, 665-685. MR2396811

[14] Dey, A.K., Martin, C.F. and Ruymgaart, F.H. (1998). Input recovery from noisy output data, using regularized inversion of Laplace transform. *IEEE Trans. Inform. Theory*, **44**, 1125–1130. MR1616699

[15] Diggle, P. J. and Hall, P. (1993). A Fourier approach to nonparametric deconvolution of a density estimate. *J. Roy. Statist. Soc. Ser.* ***B***, **55** 523–531. MR1224414

[16] Fan, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problem. *Ann. Statist.*, **19**, 1257-1272. MR1126324

[17] Fan, J. and Koo, J. (2002). Wavelet deconvolution. *IEEE Trans. Inform. Theory*, **48**, 734–747. MR1889978

[18] Gasser, T. and Müller, H-G. (1984). Estimating regression functions and their derivatives by the kernel method. *Scand. J. Statist.*, **11**, 171–185. MR0767241

[19] Gasser, T., Müller, H-G. and Mammitzsch, V. (1985). Kernels for nonparametric kernel estimation. *J. Roy. Statist. Soc. Ser.* ***B***, **47**, 238–252. MR0816088

[20] Gafni, A., Modlin, R. L. and Brand, L. (1975). Analysis of fluorescence decay curves by means of the Laplace transformation. *Biophys. J.*, **15**, 263–280.

[21] Gendre, X. (2013). Model selection and estimation of a component in additive regression. *ESAIM: Probability and Statistics*, *to appear.*

[22] Goh, V., Halligan, S., Hugill, J.A., Gartner, L. and Bartram, C.I. (2005). Quantitative colorectal cancer perfusion measurement using dynamic contrastenhanced multidetector-row computed tomography: effect of acquisition time and implications for protocols. *J. Comput. Assist. Tomogr.*, **29**, 59–63.

[23] Goh, V. and Padhani, A. R. (2007). Functional imaging of colorectal cancer angiogenesis. *Lancet Oncol.*, **8**, 245–255.

[24] Gradshtein, I.S. and Ryzhik, I.M. (1980). *Tables of Integrals, Series, and Products.* Academic Press, New York. MR0582453

[25] Gripenberg, G., Londen, S.O. and Staffans, O. (1990). *Volterra Integral and Functional Equations.* Cambridge University Press, Cambridge. MR1050319

[26] Johnstone, I.M., Kerkyacharian, G., Picard, D. and Raimondo, M. (2004). Wavelet deconvolution in a periodic setting. *J. Roy. Statist. Soc. Ser.* ***B***, **66**, 547–573 (with discussion, 627–657). MR2088290

[27] Lakowicz, J.R. (2006). *Principles of Fluorescence Spectroscopy.* Kluwer Academic, New York.

[28] Lamm, P. (1996). Approximation of ill-posed Volterra problems via predictor-corrector regularization methods. *SIAM J. Appl. Math.*, **56**, 524–541. MR1381658

[29] Laurent, B. and Massart, P. (1998). Adaptive estimation of a quadratic functional by model selection. *Ann. Statist.*, **28**, 1302–1338. MR1805785

[30] LePage, W.R. (1961). *Complex Variables and the Laplace Transform for Engineers.* Dover, New-York. MR0117514

[31] Lepski, O.V. (1991). Asymptotic mimimax adaptive estimation. I: Upper bounds. Optimally adaptive estimates. *Theory Probab. Appl.*, **36**, 654–659.

[32] Lepski, O.V., Mammen, E. and Spokoiny, V.G. (1997). Optimal spatial adaptation to inhomogeneous smoothness: an approach based on kernel estimates with variable bandwidth selectors *Ann. Statist.*, **25**, 929–947. MR1447734

[33] Lien, T.N., Trong, D.D. and Dinh, A.P.N. (2008). Laguerre polynomials and the inverse Laplace transform using discrete data *J. Math. Anal. Appl.*, **337**, 1302–1314. MR2386379

[34] Maleknejad, K., Mollapourasl, R. and Alizadeh, M. (2007). Numerical solution of Volterra type integral equation of the first kind with wavelet basis. *Appl. Math.Comput.*, **194**, 400–405. MR2385912

[35] McKinnon, A. E., Szabo, A. G. and Miller, D. R. (1977). The deconvolution of photoluminescence data. *J. Phys. Chem.*, **81**, 1564–1570.

[36] Meister, A. (2009). *Deconvolution Problems in Nonparametric Statistics (Lecture Notes in Statistics).* Springer-Verlag, Berlin. MR2768576

[37] Miles, K. A. (2003). Functional CT imaging in oncology. *Eur. Radiol.*, 13 - suppl. 5, **M134-8**.

[38] O'Connor, D. V., Ware, W. R. and Andre, J. C. (1979). Deconvolution of fluorescence decay curves. A critical comparison of techniques. *J. Phys. Chem.*, **83**, 1333–1343.

[39] Padhani, A. R. and Harvey, C. J. (2005). Angiogenesis imaging in the management of prostate cancer. *Nat. Clin. Pract. Urol.*, **2**, 596–607.

[40] Pensky, M. and Vidakovic, B. (1999). Adaptive wavelet estimator for nonparametric density deconvolution. *Ann. Statist.*, **27**, 2033–2053. MR1765627

[41] Polyanin, A.D. and Manzhirov, A.V. (1998). *Handbook of Integral Equations*, CRC Press, Boca Raton, Florida. MR1790925

[42] Stefanski, L. and Carroll, R. J. (1990). Deconvoluting kernel density estimators. *Statistics*, **21**, 169–184.

[43] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*, Springer, New York. MR2724359

[44] Weeks, W.T. (1966). Numerical inversion of Laplace transforms using Laguerre functions. *J. Assoc. Comput. Machinery*, **13**, 419–429. MR0195241