Universität Stuttgart

Fachbereich Mathematik

Deconvolving compactly supported densities

Alexander Meister

Preprint 2005/007

Universität Stuttgart
Fachbereich Mathematik

Deconvolving compactly supported densities

Alexander Meister

Preprint 2005/007

Fachbereich Mathematik
Fakultät Mathematik und Physik
Universität Stuttgart
Pfaffenwaldring 57
D-70 569 Stuttgart

E-Mail: preprints@mathematik.uni-stuttgart.de
WWW: http://www.mathematik/uni-stuttgart.de/preprints

ISSN **1613-8309**

1. Introduction

The general problem of estimating a density function based on data, which are corrupted by additive measurement error, has stimulated great research activity for both theoretical and practical matters. In literature, this topic has become known as density deconvolution. In the mathematical model, one observes the data $Y_1, \ldots, Y_n$ where $Y_j = X_j + \varepsilon_j$. The contamination of the data is represented by the random variables ε_j's with density g; while our goal is estimating the density f of the X_j's. All $X_1, \varepsilon_1, \ldots, X_n, \varepsilon_n$ are assumed to be independent. Density deconvolution can be classified as an inverse problem in the field of nonparametric statistics.

As a standard procedure, kernel methods, combined with Fourier inversion, have been proposed in Carroll & Hall (1988), Devroye (1989), Stefanski & Carroll (1990), for example. Alternative techniques involving wavelet approaches are studied in e.g. Pensky & Vidakovic (1999). For recent contributions, see e.g. Delaigle & Gijbels (2002, 2004a, 2004b), Hall & Qiu (2005), Butucea & Matias (2005), Meister (2004, 2006a, 2006b), Hall & Meister (2005).

Despite those scientific interests, density deconvolution is often criticized, firstly for the essential assumption of a known error density g, which is used in the construction of deconvolution estimators and, in addition, the Fourier transform of density g, denoted by g^{ft}, which is equivalent with the characteristic function of the corresponding random variable, is assumed to vanish nowhere; secondly, for slow rates of convergence for very smooth g.

The framework of Efromovich (1997) and Neumann (1997) is a concession to the ignorance of g, which occurs frequently in practice; but those approaches assume the availability of additional direct observations from g; that restricts the applicability to some cases where the system of measurement can be calibrated somehow. Meister (2004) shows that misspecifying g may have fatal consequences for the asymptotic quality of the deconvolution kernel estimator. The condition of perfectly known g can be relaxed, see Butucea & Matias (2005) and Meister (2006a), where semiparametric models allow only a scaling parameter or the variance of the error density to be unknown; however, in order to keep f identifiable, an additional lower bound on $|f^{ft}|$ is required. Also, the condition $g^{ft}(t) \neq 0$, which rules out important densities such as uniform densities, can be relaxed for the case where g^{ft} has isolated zeros (see Devroye (1989), Hall & Meister (2005)). But, if g^{ft} vanishes on an open and non-empty interval, then the estimation problem becomes non-identifiable, in general.

In Goldenshluger (2002), faster rates of convergence are derived for a bivariate circular deconvolution model where the density of an angle is to be estimated and g^{ft} is assumed to be exactly known and non-vanishing.

In the current note, we focus on univariate densities f which are compactly supported. That condition seems realistic in many practical applications; we refer to problems where the probability mass of the X_j's is restricted to some bounded region, due to some a-priori knowledge. Combined with usual smoothness conditions, that assumption allows us to improve the asymptotic quality of the estimation of f, with respect to both major items of criticism against deconvolution, as mentioned above. In particular, we are able to construct an estimator when only the restriction of g^{ft} to a compact interval around zero is known. Hence, unlike the semiparametric approaches to unknown g, as studied in Butucea & Matias (2005) and Meister (2006a), we are considering a comprehensive nonparametric class of densities competing to be g, without assuming that the shape of g is known up to a scaling parameter. Also, we are able to estimate f although g^{ft} may vanish on an open, non-empty set. Furthermore, our framework allows us to derive faster rates than those whose optimality has been established in Fan (1991a, 1993) for smooth f with unbounded support; and, in a special case, our rates are in the line with those derived in Goldenshluger (2002) for circular densities with known g.

We mention that the results of this paper may also have applications beyond density estimation, e.g. in the related field of image reconstruction when an image with bounded domain is to be deblurred from noise and some point-spread effects with imperfectly known distribution.

2. Methodology

In order to describe our estimation method, we need to specify the exact conditions on f and on g. With respect to the error density g, we assume knowledge of $g^{ft}(t)$ for $t \in [-\nu, \nu]$ only, and the membership of g in the density class

$$\mathcal{G}_{\mu,\nu} \;:=\; \big\{ g \text{ density } : |g^{ft}(t)| \geq \mu \,,\, \forall |t| \leq \nu \big\} \tag{1}$$

for some $\mu, \nu > 0$. Note that, whenever g is perfectly known, there are always some appropriate μ, ν so that $\mathcal{G} = \{g\}$. On the other hand, that framework allows us to consider a nonparametric class of error densities, what can be seen as follows: if g^{ft} is positive, convex and decreases on $[-\nu, \nu]$ monotonously, then Polya's criterion (see Lukacs (1970), p. 83, Theorem 4.3.1) allows g^{ft} to be continued on $[\nu, \infty)$ in lots of different ways.

Concerning the target density f, we consider densities whose support is included into a compact interval $[-S, S]$, which, in addition, satisfy some common smoothness assumptions, given by a uniform bound on a Sobolev norm. The densities are collected into the set

$$\mathcal{F}_{S;C,\beta} \;:=\; \Big\{ f \text{ density } : \int_{-S}^{S} f(x)dx \;=\; 1 \quad \text{and} \quad \int |f^{ft}(t)|^2 (1+t^2)^\beta dt \;\leq\; C \Big\} \tag{2}$$

with $S, C, \beta > 0$. Note that, for integer β, any density f supported on $[-S, S]$ which is β-fold continuously differentiable on the whole real line is contained in $\mathcal{F}_{S;C,\beta}$ for C sufficiently large.

Now we focus on constructing an appropriate estimator for f, motivated by the reconstructability of f^{ft} from the data. Due to (1), the direct empirical access to $f^{ft}(t)$, obtained by Fourier inversion, is restricted to $t \in [-\nu, \nu]$. For a bandlimiting sequence $(\omega_n)_n$ tending to infinity, the integral of $|f^{ft}(t)|^2$ on $|t| > \omega_n$ converges to zero with specific rates, due to our smoothness assumptions. Therefore, we have a gap for $|t| \in (\nu, \omega_n)$ where $f^{ft}(t)$ needs to be determined. However, condition (2) implies the existence of all moments of f or X and, hence, we may employ the Taylor expansion around 0 for characteristic functions for all t and integers $m > 0$,

$$f^{ft}(t) \;=\; \sum_{j=0}^{m} \frac{i^j}{j!} \big(EX^j\big) t^j \;+\; R_m(t), \tag{3}$$

where R_m denotes the residual term (see e.g. Lukacs (1970), Theorem 2.3.1). To get an upper bound on $|R_m(t)|$, we apply Lagrange's representation, leading to

$$|R_m(t)| \;\leq\; |t|^{m+1} \cdot E|X|^{m+1}/(m+1)! \;\leq\; |S\,t|^{m+1}/(m+1)!$$

under condition (2). Via Stirling's formula, we obtain

$$|R_m(t)| \;\leq\; O(1) \cdot \Big|\frac{S\,e\,t}{m+1}\Big|^{m+1} \cdot m^{-1/2}\,, \tag{4}$$

where $O(1)$ does not depend on t. Since $R_m(t) \to 0$, for any t, as $m \to \infty$, the function $f^{ft}(t)$ may be represented by the pointwise limit of its (complex-valued) Taylor series for all t. Hence, $f^{ft}(t)$, for all $t \in \mathbb{R}$, is uniquely determined by its restriction to $|t| \leq \nu$, from what follows that the empirically gained information about $f^{ft}(t)$ on the domain $t \in [-\nu, \nu]$ makes $f^{ft}(t)$ accessible even on the whole interval $t \in [-\omega_n, \omega_n]$. That ensures identifiability of f in the underlying estimation problem.

That essential representability of $f^{ft}(t)$ by its Taylor expansion, for all t, inspires us to employ a polynomial approach to estimate f^{ft}. The projection of the empirical Fourier transform of f onto the space of all polynomials with degree $\leq m_n$ on the domain $[-\nu, \nu]$ is given by

$$\hat{\Psi}_n(t) \;=\; \frac{1}{n} \sum_{j=1}^{n} \sum_{k=0}^{m_n} P_k(t) \int_{-\nu}^{\nu} P_k(s) \exp(isY_j)/g^{ft}(s) ds, \tag{5}$$

where m_n is a positive integer, $P_k(t) = \big((2k+1)/(2\nu)\big)^{1/2}\,\tilde{P}_k(t/\nu)$ and $\tilde{P}_k$ denotes the kth Legendre polynomial on $[-1,1]$. We refer to a widely-used equality for Legendre polynomials

$$\int_{-1}^{1} |\tilde{P}_k(x)|^2 dx \;=\; 2/(2k+1)\,. \tag{6}$$

Therefore, P_k, integer $k \geq 0$, are orthonormal with respect to the inner product of $L_2([-\nu,\nu])$. Considering that $\hat{\Psi}_n(t)$ is a polynomial on $t \in [-\nu,\nu]$, its domain may easily be continued onto the whole real line in a natural way. So $\hat{\Psi}_n(t)$ is well-defined by (5) for all real t. Fourier inversion of (5) leads to the following density estimator of f,

$$\hat{f}_n(x) \;=\; \mathrm{Re}\,\frac{1}{2\pi}\int L(t/\omega_n)\exp(-itx)\hat{\Psi}_n(t)dt, \tag{7}$$

where Re denotes the real part. Throughout this note, we assume that the function $L(s)$ is supported on $[-1,1]$; also, $|L(s)| \leq 1$, for all s; $L(0) = 1$ and $L(s)$ is continuous at $s = 0$. So L can be viewed at as the Fourier transform of an appropriate kernel function. The parameters m_n and ω_n remain to be selected with respect to the sample size n.

3. Asymptotic properties

In this section, we focus on the asymptotic behavior of the MISE (mean integrated squared error) of our estimator (7), given by

$$\mathrm{MISE}_n(g,f) \;:=\; E\|\hat{f}_n - f\|^2,$$

where $\|\cdot\|$ denotes the $L_2(\mathbb{R})$-norm throughout this paper. For the sake of generality, we allow the endpoint S to increase in the sample size n. That relaxes the condition of a strict concentration of the probability mass and will have some effects on the rates. Therefore, we have to assume that, at least, some $\tilde{S}_n$ is known with

$$1 \;\leq\; \tilde{S}_n/S_n \;\leq\; O(1)\,. \tag{8}$$

The following Theorem 1(a) gives the rates of convergence when considering $\mathrm{MISE}_n(g,f)$ uniformly on the classes defined for g and f in (1) and (2), respectively. Part (b) of the theorem establishes consistency in a general setting; more concretely, in absence of smoothness assumptions.

Theorem 1 *The restriction of $g^{ft}(t)$ to $t \in [-\nu,\nu]$ as well as $\tilde{S}_n \leq O\big((\ln n)^\delta\big)$, $\delta \in [0,1)$, as in (2) are assumed to be known. We select $m_n = \lfloor C_1 \ln n/(\ln\ln n)\rfloor$ and $\omega_n = C_2\tilde{S}_n^{-1}\cdot m_n$ with some constants $C_1 \in \big(0,1/2\big)$ and $C_2 \in \big(0,1/(2\,e)\big)$. Then,*
(a) if L satisfies $|L(s)-1| = O(|s|^\beta)$, in addition, we obtain

$$\sup_{g\in\mathcal{G}_{\mu,\nu}}\;\sup_{f\in\mathcal{F}_{S_n;C,\beta}} MISE_n(g,f) \;=\; O\Big(\big(\ln n\big)^{-2\beta(1-\delta)}(\ln\ln n)^{2\beta}\Big)\,,$$

(b) for any density f supported on fixed $[-S,S]$, we have

$$MISE_n(g,f) \overset{n\to\infty}{\longrightarrow} 0 \quad , \qquad \forall g \in \mathcal{G}_{\mu,\nu}\,.$$

Note that the condition on L in (a), along with the previously assumed restrictions for L, are satisfied by e.g. $L(s) = \chi_{[-1,1]}(s)$, where χ_I denotes the indicator function of a set I. As an example, we consider the case where g is a known normal density. Then we may consider the SDDKE (standard deconvolution density kernel estimator) as a competing method. We recall its definition (e.g. Carroll & Hall (1988), Stefanski & Carroll (1990)),

$$\hat{f}_{n,K} \;=\; \frac{1}{2\pi n}\int K^{ft}(t/\omega_n)\sum_{j=1}^{n}\exp\big(it(Y_j - x)\big)\,/\,g^{ft}(t)\,dt\,,$$

with a kernel function K; as usual, we restrict our consideration to those kernels whose Fourier transforms K^{ft} are compactly supported and bounded away from zero in a neighborhood of $t = 0$. So K^{ft} corresponds to function L in estimator (7). The SDDKE has been considered for smooth f with unbounded support; i.e., corresponding to (2) when the assumed restriction of the probability mass to $[-S_n, S_n]$ is removed and only the Sobolev condition involving parameter β is required. Then, the SDDKE achieves the optimal rate $O\big((\ln n)^{-\beta}\big)$, what has been derived in Fan (1991a, 1993) for Hölder classes and extended to Sobolev classes (see Neumann (1997), Hesse & Meister (2004)). However, we notice the surprising fact that the rate derived in Theorem 1(a) is faster if $\delta < 1/2$. Obviously, in that case, the additional condition of a compact support allows us to improve the rate. The rates in the case of fixed support, i.e. $\delta = 0$, correspond to those derived in Goldenshluger (2002) for circular densities of an angle (so also with fixed support $(0, 2\pi)$), where g is known. We see that we have to pay for $\tilde{S}_n \to \infty$ with respect to the rates. Nevertheless, the following proposition shows that, unlike estimator (7), the SDDKE is unable to take advantage of the bounded support of f and provides slower rates, compared to estimator (7) in the case $\delta < 1/2$.

Proposition 1 *Assume that g is a normal density; $S_n = S$; and C sufficiently large. Then, there is a constant $c > 0$ so that*

$$\sup_{f\in\mathcal{F}_{S;C,\beta}} E\,\|\hat{f}_{n,K} - f\|^2 \;\geq\; c\cdot(\ln n)^{-\beta}\,.$$

Furthermore, if only $g \in \mathcal{G}_{\mu,\nu}$ is known, we still obtain the rate given in Theorem 1(a) for estimator (7); while the SDDKE cannot be used as it requires full knowledge of g in its construction. That refers to the case where g is imperfectly known, i.e. we only have the information $g \in \mathcal{G}_{\mu,\nu}$, as well as to those g, where g^{ft} vanishes outside a compact interval around zero. An example for such a density is given in Section 4. Therefore, similar to Goldenshluger (2002), we are able to keep those rates even for error densities which are smoother than supersmooth, in the terminology of Fan (1991a, 1993). Also, the SDDKE is very sensitive when plugging-in an incorrect error density; in particular, with respect to the decay of g^{ft}, see Meister (2004). On the other hand, estimator (7) does not care about some misspecification of $g^{ft}(t)$ as long as $t \in \mathbb{R}\backslash[-\nu, \nu]$. Therefore, estimator (7) is more robust with respect to g.

In the case where g is known, we learn from Theorem 1(b) that one is able to estimate any compactly supported density consistently, if bounds on the support boundaries are known, but without any conditions on g. On the other hand, in the case of densities with unbounded support, at least some conditions referring to the set $\{t : g^{ft}(t) = 0\}$ are required, see Devroye (1989) and Meister (2006b). Also, note that, as we are considering densities with a uniformly bounded support, the $L_2(\mathbb{R})$-distance of densities dominates the $L_1(\mathbb{R})$-distance. So if we truncate estimator (7) to a bounded set $[-S', S']$ with $S' \geq S$, consistency of that truncated estimator, with respect to $L_1(\mathbb{R})$, follows immediately from the convergence of the MISE.

Also, we mention that the choice of ω_n and m_n in Theorem 1 does not require any information about the smoothness of f; nor about the exact endpoints S_n, only (8) has to be satisfied.

Now we aim to establish a lower bound for the rates of convergence under the assumptions of Theorem 1(a), referring to any estimator. The proof requires a new concept for densities with bounded support.

Theorem 2 *Under the conditions of Theorem 1(a), let $\tilde{f}_n$ be an arbitrary estimator of f based on knowledge of $g^{ft}(t)$, $t \in [-\nu, \nu]$, of $c_2 \geq S_n(\ln n)^{-\delta} \geq c_1$ for some $c_2 \geq c_1 > 0$; and the data $Y_1, \ldots, Y_n$. Suppose C to be large enough; $g^{ft}(t)$ shall be piecewise differentiable on $t \in [-\nu, \nu]$ and $g^{ft\prime}(t)$ shall be bounded on $t \in [-\nu, \nu]$; $\beta > (\delta + 1/2)/(1 - \delta)$; and $\mu \in (0, 1)$. Then, there is a constant $c > 0$ so that, for n sufficiently large, we have*

$$\sup_{g\in\mathcal{G}_{\mu,\nu}} \sup_{f\in\mathcal{F}_{S_n;C,\beta}} E\|\tilde{f}_n - f\|^2 \geq c\cdot\big(\ln n\big)^{-2\beta(1-\delta)}\cdot(\ln\ln n)^{2\beta}\,.$$

Hence, we have shown that, under slight additional technical assumptions, estimator (7) enjoys optimal rates of convergence with respect to its MISE in the underlying statistical experiment.

Finally, we mention that estimator (7) still achieves the rates given in Theorem 1(a) when the restriction of the support to $[-S,S]$ in (2) is replaced by conditions on the decay of f_X while the support of f_X may be unbounded. For example, consider densities

$$f_X(x) \leq C\exp(-d|x|^{1/\delta}) \quad , \qquad \forall|x|>1\,, \tag{9}$$

combined with the Sobolev condition. Then, we have $E|X|^l \leq (\text{const.}\cdot l)^{\delta l}$, for all integers $l>0$. Considering the specific choice of m_n and $\tilde{S}_n$ in Theorem 1(a), we receive (4) where only constant C_2 needs to be adapted; and that is all needed for the proof of the upper bound. Hence, we obtain the same rates for (9) as in Theorem 1(a). A paricularly interesting case is $\delta=1/2$, as normal densities are included. There, we notice that the SDDKE and estimator (7) provide identical rates (while use of the SDDKE necessitates full knowledge of g – as said before).

4. Simulations

In this section, we simulate data sets for our estimation problems and examinate the practical performance of our estimator (7). As the optimal rates are still logarithmic (with improved power of the logarithm), large sample sizes are required to have the estimator work well. We base our estimators on the observation of $n=1000$ independent data.

Another difficult question concerns the selection of ω_n and m_n for finite sample sizes. In Theorem 1, the parameter choice is motivated by the asymptotic quality of the estimator. However, we can derive some guidelines for the rough locations of the parameters from the theorem. For the simulations, we choose $m_n=2$, $\omega_n=1.1$. We plot three replications for each simulation. The target density f is shown as a dashed curve. Figure 1,3,5,6 are based on a unimodal density f with support $[-4,4]$; while we have a bimodal density f on $[-6,6]$ in Figure 2,4.

As function L in (7), we utilize $L(s)=\big[1-\exp(1-s^{-2})\big]\cdot\chi_{[-1,1]}(s)$, which satisfies the conditions on L in Theorem 1(a). In the light of that theoretical aspect, we could also choose the rather simple function $L=\chi_{[-1,1]}$; however, in practice, it is not favorable since the jump of $L(s)$ at $|s|=1$ may cause the estimator to oscillate too heavily; that problem is also referred to as Gibbs phenomenon.

In Figure 1 and 2, we consider the case where the error density is equal to $g_1 = g_0 * N(0,1)$; where $*$ denotes convolution; $N(0,1)$ is the standard normal density; and $g_0(x) = \big[1-\cos(x)\big]/(\pi x^2)$, having the compactly supported Fourier transform $g_0^{ft}(t)=(1-|t|)\cdot\chi_{[-1,1]}(t)$. So there is a necessity to select $\nu<1$; we choose $\nu=1/2$, i.e. we assume knowledge of $g_1^{ft}(t)$ only on its restriction to $t\in[-\nu,\nu]$.

In Figure 3 and 4, we consider the discrete distribution G of a random variable Z with $P(Z=0)=1/2$, $P(Z=(2k-1)\pi)=2/[(2k-1)^2\pi^2]$, for any integer k. As error density, we take $g_2=G*N(0,1)$. Note that $g_2^{ft}(t)$ and $g_1^{ft}(t)$ are equal to each other on their restriction to $t\in[-1,1]$. Hence, we do not change the construction of estimator (7), with respect to the case considered in Figure 1 and 2. That emphasizes the applicability of our estimator for problems with imperfectly known error distributions.

In Figure 5, we employ estimator (7) in the case of standard normal noise. There, estimator (7) faces the competition with the SDDKE (with $K^{ft}=L$, $\omega_n=1.1$), whose outcome is shown in Figure 6.

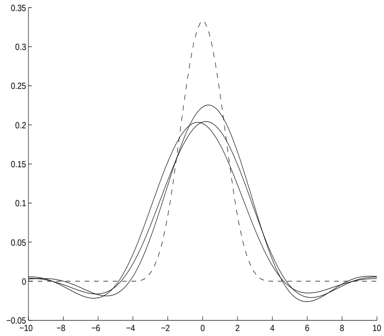

Figure 1

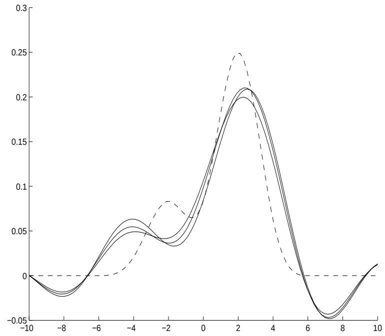

Figure 2

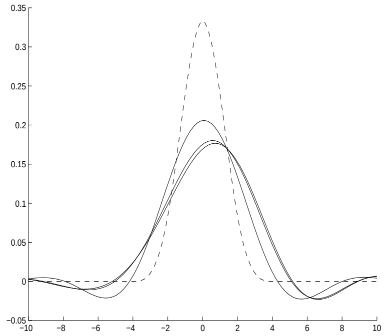

Figure 3

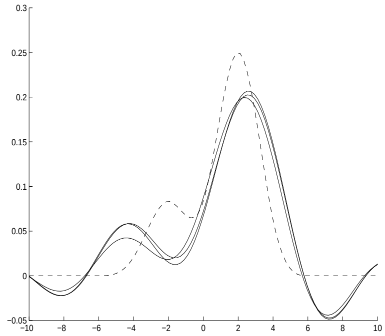

Figure 4

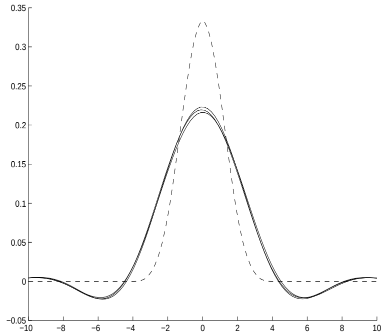

Figure 5

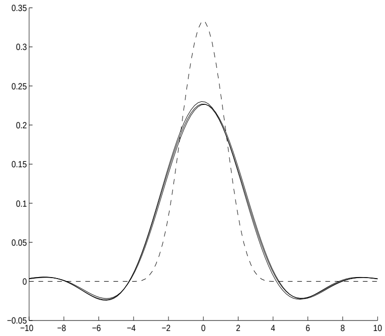

Figure 6

5. Lemmas and Proofs

Proof of Theorem 1: By using the Parseval identity, Fubini's theorem and the usual splitting into variance and bias term, we see that $\mathrm{MISE}_n(g,f)$ is bounded above by the sum of

$$\begin{aligned} V_n &:= \frac{1}{2\pi}\int_{-\omega_n}^{\omega_n} |L(t/\omega_n)|^2 \ \operatorname{var} \hat{\Psi}_n(t)\,dt\,, \\ B_{n,1} &:= \frac{1}{\pi}\int_{-\omega_n}^{\omega_n} |L(t/\omega_n)|^2 \left|E\hat{\Psi}_n(t) - f^{ft}(t)\right|^2 dt\,, \\ B_{n,2} &:= \frac{1}{\pi}\int |L(t/\omega_n) - 1|^2 |f^{ft}(t)|^2\,dt\,. \end{aligned}$$

Except for term $B_{n,2}$, the following consideration refers to part (a) as well as part (b) of the theorem. In the view of (a), the results are to be considered uniform on the classes $\mathcal{G}_{\mu,\nu}$ and

$\mathcal{F}_{S_n;C,\beta}$ while we can take an arbitrary f supported on some fixed $[-S_n, S_n]$ (so $S_n = S$) and $g \in \mathcal{G}_{\mu,\nu}$ for part (b).

With respect to V_n, we obtain

$$\begin{aligned} \operatorname{var} \hat{\Psi}_n(t) \;&\leq\; n^{-1}\, E\Big|\sum_{k}^{m_n} P_k(t) \int_{-\nu}^{\nu} P_k(s)\, \exp(isY_1)/g^{ft}(s)\, ds\,\Big|^2 \\ &\leq\; (2\nu/n) \cdot \int_{-\nu}^{\nu} \Big|\sum_{k}^{m_n} P_k(t) P_k(s)\Big|^2 |g^{ft}(s)|^{-2} ds\,, \end{aligned} \tag{10}$$

where we have employed Jensen's inequality with respect to $(2\nu)^{-1}\int_{-\nu}^{\nu}\cdots ds$. Due to (1), we have $|g^{ft}(s)|^{-2} \leq \mu^{-2}$ for all $s \in [-\nu, \nu]$. So (10) has the upper bound

$$\begin{aligned} O\big(n^{-1}\big) \int_{-\nu}^{\nu} \Big|\sum_{k}^{m_n} P_k(t) P_k(s)\Big|^2 ds \;&=\; O(n^{-1}) \cdot \sum_{k,k'}^{m_n} P_k(t) P_{k'}(t) \int_{-\nu}^{\nu} P_k(s) P_{k'}(s) ds \\ &\leq\; O\big(n^{-1}\big) \cdot \sum_{k}^{m_n} |P_k(t)|^2\,, \end{aligned}$$

due to the orthonormality of the P_k. So, for V_n, we derive

$$V_n \;\leq\; O\big(n^{-1}\big) \cdot \sum_{k}^{m_n} \int_{-\omega_n}^{\omega_n} |P_k(t)|^2 dt\,, \tag{11}$$

considering that $|L(t/\omega_n)| \leq 1$, $\forall t$. Based on the following inequality for Legendre polynomials,

$$|\tilde{P}_k(t)| \;\leq\; 2^{-k} \sum_{l=0}^{k} \binom{k}{l}^2 |t-1|^{k-l} |t+1|^l \;\leq\; 2^{-k} \binom{2k}{k} \cdot (|t|+1)^k \;\leq\; (2|t|+2)^k\,,$$

for all real t (see Koepf (1998)), we derive

$$\int_{-\omega_n}^{\omega_n} |P_k(t)|^2 dt \;=\; O\Big(\big(2 + 2\omega_n/\nu\big)^{2k+1}\Big)\,. \tag{12}$$

Inserting that into (11) gives us

$$V_n \;=\; O\Big(n^{-1}\big(2 + 2\omega_n/\nu\big)^{2m_n+1}\Big)\,. \tag{13}$$

We focus on term $B_{n,1}$. By calculating the expectation, we obtain

$$B_{n,1} \;=\; \frac{1}{\pi} \int_{-\omega_n}^{\omega_n} |L(t/\omega_n)|^2 \,\Big|\sum_{k}^{m_n} P_k(t) \int_{-\nu}^{\nu} f^{ft}(s) P_k(s) ds \;-\; f^{ft}(t)\Big|^2 dt\,. \tag{14}$$

We replace f^{ft} by the representation (3) with $m = m_n$. Since P_k, $k = 0, \ldots, m_n$, is an orthonormal base of the space $\mathcal{P}_{m_n}$ consisting of all polynomials with domain $[-\nu, \nu]$, degree $\leq m_n$ and complex coefficients, we have, for any $p \in \mathcal{P}_{m_n}$,

$$\sum_{k}^{m_n} P_k(t) \int_{-\nu}^{\nu} p(s) P_k(s) ds \;=\; p(t) \tag{15}$$

on $t \in [-\nu, \nu]$. As we have polynomials on the left as well as on the right side in (15), this equality even holds for all $t \in \mathbb{R}$. We learn from there that the polynomial part of $f^{ft}(t)$ in (3) annuls itself in (14) and we obtain (again considering that $|L(t/\omega_n)| \leq 1$, $\forall t$)

$$
\begin{aligned}
B_{n,1} \;\le\; & \frac{2}{\pi}\int_{-\omega_n}^{\omega_n} \big|R_{m_n}(t)\big|^2 dt \,+\, \frac{2}{\pi}\int_{-\omega_n}^{\omega_n} \Big|\sum_k^{m_n} P_k(t)\int_{-\nu}^{\nu} R_{m_n}(s)P_k(s)ds\Big|^2 dt \\
\le\; & O(1)\cdot\int_{-\omega_n}^{\omega_n} \big|R_{m_n}(t)\big|^2 dt \,+\, O(1)\cdot\sum_{k,k'}^{m_n}\Big(\int_{-\omega_n}^{\omega_n} P_k(t)P_{k'}(t)dt\Big)\cdot\Big(\int_{-\nu}^{\nu} R_{m_n}(s)P_k(s)ds\Big) \\
& \qquad\qquad \cdot\Big(\int_{-\nu}^{\nu} R_{m_n}(s)P_{k'}(s)ds\Big)\,.
\end{aligned} \tag{16}
$$

We apply the Cauchy-Schwarz-inequality and the orthonormality of the P_k to the second addend in (16); then we use (12) again. That gives us the inequality

$$
\begin{aligned}
B_{n,1} \;\le\; & O(1)\cdot\int_{-\omega_n}^{\omega_n} \big|R_{m_n}(t)\big|^2 dt \,+\, O(1)\cdot\int_{-\nu}^{\nu}|R_{m_n}(s)|^2 ds\cdot\Big[\sum_k^{m_n}\Big(\int_{-\omega_n}^{\omega_n}|P_k(t)|^2dt\Big)^{1/2}\Big]^2 \\
\le\; & O(1)\cdot\int_{-\omega_n}^{\omega_n} \big|R_{m_n}(t)\big|^2 dt \,+\, O\Big(m_n^2\cdot\big(2+2\omega_n/\nu\big)^{2m_n+1}\Big)\cdot\int_{-\nu}^{\nu}|R_{m_n}(s)|^2 ds\,.
\end{aligned} \tag{17}
$$

Applying (4) gives us

$$
B_{n,1} \;=\; O\Big(\omega_n^3 m_n^{-3}(S_n\, e\,\omega_n/m_n)^{2m_n}\,,\,\omega_n m_n^{-1}\cdot[2S_n\, e\,(\nu+\omega_n)/m_n]^{2m_n}\Big)\,. \tag{18}
$$

Due to $\omega_n/m_n = C_2/S_n$, following from the selection rules in the theorem, the latter term in (18) dominates the first one, implying

$$
B_{n,1} \;=\; O\Big(\big[2\,S_n\, e\,(\nu+\omega_n)/m_n\big]^{2m_n}\Big)\,. \tag{19}
$$

Term $B_{n,2}$ remains to be studied. In part (b), we use the fact that any density with a bounded support has a finite $L_2(\mathbb{R})$-norm, from what follows $B_{n,2}\to 0$ by the Parseval identity, the boundedness of $L(s)$ and its continuity at $s=0$, since $\omega_n\to\infty$. However, the rates are unspecified in that framework. The specific choice of ω_n and m_n as stated guarantees convergence of both (13) and (19) to zero, which proves Theorem 1(b). In part (a), the Sobolev condition contained in (2) gives us $B_{n,2} = O\big(\omega_n^{-2\beta}\big)$, due to the additional assumption on L in (a); note that this term is the same as in common density estimation in Sobolev classes. We summarise all terms which are significant for the upper bound,

$$
\begin{aligned}
& \sup_{g\in\mathcal{G}_{\mu,\nu}}\ \sup_{f\in\mathcal{F}_{S_n;C,\beta}}\ \mathrm{MISE}_n(g,f) \\
& \qquad\qquad = \; O\Big(\omega_n^{-2\beta}\,,\, n^{-1}\big(2+2\omega_n/\nu\big)^{2m_n+1}\,,\,\big[2S_n\, e\,(\nu+\omega_n)\,/\,m_n\big]^{2m_n}\Big),
\end{aligned}
$$

Inserting m_n and ω_n as in the theorem gives the rate of convergence as stated. ■

Proof of Proposition 1: Take φ_b from the proof of Theorem 2. We consider the density

$$
\kappa_n(x) \;=\; \delta_n\cdot\varphi_{\gamma_n}(x) \,+\, (1-\delta_n)\cdot\varphi_S(x)\,,
$$

where we set $\delta_n = \text{const.}\cdot\gamma_n^{-\beta-1/2}$ with an appropriate constant; where $\gamma_n\uparrow\infty$. Then, $\kappa_n\in\mathcal{F}_{S;C,\beta}$ is guaranteed for C large enough. The variance of the SDDKE is equal to

$$[\omega_n/(2\pi n)] \cdot \int_{-1}^{1} \left|K^{ft}(s)\right|^2 \Big(\left|g^{ft}(s\omega_n)\right|^{-2} - |f^{ft}(s\omega_n)|^2\Big)\, ds\,;$$

so it has the lower bound const. $\cdot\, n^{-1}\omega_n \exp(d\omega_n^2)$ for some $d > 0$.

The bias term is as $\int_{\omega_n}^{\infty} \left|\kappa_n^{ft}(t)\right|^2 dt$ due to the compact support of K^{ft}. As we can neglect the term involving φ_S^{ft}, based on (26), we may set $\gamma_n = $ const. $\cdot\, \omega_n$, and we obtain const. $\cdot\, \omega_n^{-2\beta}$ as a lower bound for the bias.

Hence, the optimal choice of ω_n within that framework is $\omega_n = $ const. $\cdot\, (\ln n)^{1/2}$, leading to the rate given in the proposition. ■

As an important tool for the proof of Theorem 2, we need some technical properties of Legendre polynomials, which are given in the following lemma.

Lemma 1 *There are constants $c_0, c_1, c_2 > 0$, $\tau \in (0,1)$ so that we have more than $c_1 k$ disjoint intervals with the length $\geq c_0 k^{-1}$, which are included into $(-\tau,\tau)$, and $|\tilde{P}_k(x)|^2 > c_2 k^{-1}$ holds for all x contained in one of those intervals, for all $k \geq K$ with K sufficiently large.*

Proof of Lemma 1: We utilize the following important inequality for Legendre polynomials

$$|\tilde{P}_k(\cos\theta)|^2 \;\leq\; 2/(\pi\, k\, \sin\theta) \quad , \quad \theta \in [0,\pi]\,. \tag{20}$$

From there, we derive

$$\int_{-1}^{-\tau} |\tilde{P}_k(x)|^2 dx \;\leq\; 2[\pi - \arccos(-\tau)]/(\pi k)\;, \quad \int_{\tau}^{1} |\tilde{P}_k(x)|^2 dx \;\leq\; 2\arccos\tau/(\pi k)\,,$$

where $\arccos x \in [0,\pi]$, for any $x \in [-1,1]$. Therefore, considering (6), one is able to select $\tau < 1$ sufficiently close to 1 so that

$$\int_{-\tau}^{\tau} |\tilde{P}_k(x)|^2 dx \;\geq\; c_0 k^{-1} \quad , \quad \text{constant } c_0 > 0\,. \tag{21}$$

Now we consider the set

$$G_k(c) \;=\; \left\{x \in (-\tau,\tau)\,:\, |\tilde{P}_k(x)|^2 > ck^{-1}\right\}.$$

For all $x \in [-\tau,\tau]$, we may establish

$$|\tilde{P}_k(x)|^2 \;\leq\; c_3 k^{-1} \quad , \quad \text{constant } c_3 > 0\,, \tag{22}$$

following from (20). We derive by (21) and (22),

$$\begin{aligned} c_0 k^{-1} \;&\leq\; \int_{G_k(c)} |\tilde{P}_k(x)|^2 dx \;+\; \int_{[-\tau,\tau]\backslash G_k(c)} |\tilde{P}_k(x)|^2 dx \\ &\leq\; c_3 k^{-1}\mu(G_k(c)) \;+\; ck^{-1}\mu\big([-\tau,\tau]\backslash G_k(c)\big) \;\leq\; c_3 k^{-1}\mu(G_k(c)) \;+\; 2\tau c k^{-1}\,, \end{aligned}$$

where μ denotes the Lebesgue measure. Therefore, we have $\mu(G_k(c)) \;\geq\; (c_0 - 2\tau c)/c_3$. Under suitable selection of c, we obtain

$$\mu(G_k(c)) \;\geq\; \text{const.} \;>\; 0 \quad , \quad \forall k\,. \tag{23}$$

We set c_2, as it occurs in the lemma, equal to that $c > 0$ satisfying (23). Furthermore, we use some results about the zeros of the Legendre polynomials: $\tilde{P}_k$ possesses k different zeros, which all lie in $(-1,1)$; so they are representable by $x_j = \cos\theta_j$, $0 < \theta_1 < \cdots < \theta_k < \pi$. We use the inequality of Bruns (see Szegö (1936)), saying that

$$(j-1/2)\pi/(k+1/2) \, < \, \theta_j \, < \, j\pi/(k+1/2) \quad , \quad \forall 1 \le j \le k \, .$$

It follows from there that those zeros x_j which are contained in $(-\tau,\tau)$, $\tau < 1$, satisfy

$$c_4 k^{-1} \, < \, x_j - x_{j+1} \, < \, c_5 k^{-1} \, . \tag{24}$$

Also, for large k, we notice that there are zeros $x_l, x_{l'}$ with $x_l < -\tau < \tau < x_{l'}$. Since $|\tilde{P}_k|$ has at most $k-1$ local maxima (with $\tilde{P}_k'(x) = 0$) and, on the other hand, there is at least one maximum of $|\tilde{P}_k|$ between each x_{j+1} and x_j, we have exactly one maximum in (x_{j+1}, x_j), denoted by y_j; and we notice that $|\tilde{P}_k|$ increases on (x_{j+1}, y_j) and decreases on (y_j, x_j). From there, we receive representability of $G_k(c)$ by the union of finitely many disjoint intervals denoted by $I_{j,k}(c)$,

$$G_k(c) \, = \, \bigcup_{j \in J_k} I_{j,k}(c) \quad , \qquad I_{j,k}(c) \subseteq (x_{j+1}, x_j) \subseteq (-\tau,\tau) \, ,$$

with an appropriate set J_k of integers, which contains less than k elements. So we have

$$\sum_{j \in J_k} \mu\big(I_{j,k}(c)\big) \, = \, \mu\big(G_k(c)\big) \, \ge \, \text{const.} \, > 0 \, ,$$

due to (23). For further consideration, we introduce the collection

$$J_k'(c,d) \, = \, \big\{ j \in J_k \, : \, \mu(I_{j,k}(c)) \ge d k^{-1} \big\} \, ,$$

and we derive the inequality

$$\begin{aligned} 0 \, < \, \text{const.} \, &\le \sum_{j \in J_k'(c,d)} \mu(I_{j,k}(c)) \, + \sum_{j \in J_k \setminus J_k'(c,d)} \mu(I_{j,k}(c)) \\ &\le \, d k^{-1} k \, + \, c_5 k^{-1} \cdot \# J_k'(c,d) \, , \end{aligned}$$

due to (24). Therefore, we have

$$\# J_k'(c,d) \, \ge \, (\text{const.} - d) c_5^{-1} \cdot k \, ,$$

so we see that, by choosing $d > 0$ suitably, we obtain $\# J_k'(c,d) \, > \, c_1 \cdot k$ with a constant $c_1 > 0$; furthermore we set $c_0 = d$ so that all constants c_1, c_0, c_2 occurring in the lemma have been defined. ■

Proof of Theorem 2: We introduce the stretched and truncated Legendre polynomials $Q_k(x) = \tilde{P}_k(2x/S_n) \cdot \chi_{[-S_n/2, S_n/2]}(x)$. Also, we specify the density $\varphi(x) = a \cdot \exp\big(-1/(1-x^2)\big) \cdot \chi_{[-1,1]}(x)$ with the appropriate constant $a > 0$. We set $\varphi_b(x) = b\varphi(bx)$ and $f = \varphi_{1/S_n}$. Furthermore, we define

$$f_n(x) \, = \, f(x) \, + \, \alpha_n \cdot \big(Q_{K_n} * \varphi_{b_n}\big)(x) \, ,$$

where $*$ denotes convolution; $\alpha_n \downarrow 0$, $b_n \uparrow \infty$, $K_n \uparrow \infty$ are still to be selected. That choice of the parameters, combined with $\beta > (\delta + 1/2)/(1-\delta)$, guarantees that f, f_n are densities which are supported on $[-S_n, S_n]$, for n sufficiently large. In order to verify the membership of f, f_n in $\mathcal{F}_{S_n;C,\beta}$, we have to check the Sobolev condition. Function f, which does not depend on n, is differentiable infinitely often where f along with all of its derivatives is compactly supported; this implies

$$\int |f^{ft}(t)|^2 (1+t^2)^\beta \, dt \, \le \, C \, ,$$

for C large enough. Concerning f_n, we need to check, in addition,

$$\begin{aligned}
\alpha_n^2 \int |Q_{K_n}^{ft}(t)|^2 |\varphi^{ft}(t/b_n)|^2 (1+t^2)^\beta \, dt \; &\leq \; O\big(\alpha_n^2 b_n^{2\beta+1}\big) \int |Q_{K_n}^{ft}(b_n s)|^2 |\varphi^{ft}(s)|^2 (1+s^2)^\beta \, ds \\
&\leq \; O\big(\alpha_n^2 b_n^{2\beta}\big) \int |Q_{K_n}^{ft}(t)|^2 \, dt \cdot \sup_s |\varphi^{ft}(s)|^2 \, (1+s^2)^\beta \\
&\leq \; O\big(\alpha_n^2 b_n^{2\beta}\big) \int_{-S_n/2}^{S_n/2} |Q_{K_n}(t)|^2 \, dt \; \leq \; O\big(\alpha_n^2 b_n^{2\beta} S_n\big) \int_{-1}^{1} |\tilde{P}_{K_n}(t)|^2 \, dt \\
&\leq \; O\big(\alpha_n^2 b_n^{2\beta} K_n^{-1} S_n\big) \, , \qquad (25)
\end{aligned}$$

where we have used that $|\varphi^{ft}(s) s^m|$, for any integer $m > 0$, can be bounded above by partial integration, as follows,

$$|\varphi^{ft}(s) s^m| \; = \; \Big| \int \Big(\frac{d^m}{dt^m} \exp(its) \Big) \varphi(t) dt \Big| \; = \; \Big| \int \exp(its) \varphi^{(m)}(t) dt \Big| \; \leq \; \int |\varphi^{(m)}(t)| dt \, , \qquad (26)$$

from what we obtain $|\varphi^{ft}(s)| \leq C_m \cdot |s|^{-m}$, $\forall s$, for some $C_m > 0$; combined with $|\varphi^{ft}(s)| \leq 1$, as φ is a density, we choose $m > \beta$ and, therefore, we have a uniform bound for $|\varphi^{ft}(s)|^2 (1+s^2)^\beta$ on $s \in \mathbb{R}$. Also, (6) has been employed in (25). The Sobolev norm is bounded with respect to n when setting

$$\alpha_n \; = \; \text{const.} \cdot b_n^{-\beta} K_n^{1/2} S_n^{-1/2} \, . \qquad (27)$$

with an appropriate positive constant. Under this selection, we get $f, f_n \in \mathcal{F}_{S_n;C,\beta}$ for C large enough.

We choose the error density

$$\tilde{g}(x) \; = \; \big[1 - \cos(\nu' x)\big] / (\pi \nu' x^2) \quad , \quad \nu' > 0 \, ,$$

having the Fourier transform $\tilde{g}^{ft}(t) \; = \; (1 - |t/\nu'|) \cdot \chi_{[-\nu',\nu']}(t)$. By choosing ν' sufficiently large with respect to some arbitrary $\mu < 1, \nu > 0$, we may verify $\tilde{g} \in \mathcal{G}_{\mu,\nu}$.

Now, suppose an arbitrary estimator $\tilde{f}_n$. Then, we have

$$\begin{aligned}
\sup_{g \in \mathcal{G}_{\mu,\nu}} \; \sup_{f \in \mathcal{F}_{S_n;C,\beta}} \; E_{\tilde{g},f} \|\tilde{f}_n - f\|^2 \; &\geq \; (1/2) \cdot E_{\tilde{g},f} \|\tilde{f}_n - f\|^2 \; + \; (1/2) \cdot E_{\tilde{g},f_n} \|\tilde{f}_n - f_n\|^2 \\
&\geq \; (1/4) \cdot \|f - f_n\|^2 \cdot \int \cdots \int \prod_{j=1}^n \min\big\{ (f * \tilde{g})(x_j), (f_n * \tilde{g})(x_j) \big\} \, dx_1 \cdots dx_n \\
&\geq \; (1/4) \cdot \|f - f_n\|^2 \cdot \Big[1 - \int \big| (f * \tilde{g})(x) - (f_n * \tilde{g})(x) \big| dx \Big]^n \, .
\end{aligned}$$

Hence, the supremum of the MISE is bounded below by $c \cdot \|f - f_n\|^2$, constant $c > 0$, if

$$\int \big| (f * \tilde{g})(x) - (f_n * \tilde{g})(x) \big| dx \; \leq \; O(1/n) \qquad (28)$$

is valid. In order to verify (28), we introduce the Cauchy density $f_0(x) = \big[\pi(1+x^2)\big]^{-1}$; employing the Cauchy-Schwarz-inequality leads to

$$\int \big|(f * \tilde g)(x) - (f_n * \tilde g)(x)\big| dx \;\le\; \Big(\int \big|[(f-f_n)*\tilde g](x)\big|^2 \,/\, f_0(x)\, dx\Big)^{1/2}$$
$$\le\; O(1)\cdot \Big(\int \big|[(f-f_n)*\tilde g](x)\big|^2 \cdot (1+x^2) dx\Big)^{1/2}$$
$$\le\; O(1)\cdot \Big(\int \big|f^{ft}(t) - f_n^{ft}(t)\big|^2 \cdot \big|\tilde g^{ft}(t)\big|^2 dt \,+\, \int \big|f^{ft\prime}(t) - f_n^{ft\prime}(t)\big|^2 \cdot \big|\tilde g^{ft}(t)\big|^2 dt$$
$$+\, \int \big|f^{ft}(t) - f_n^{ft}(t)\big|^2 \cdot \big|\tilde g^{ft\prime}(t)\big|^2 dt\Big)^{1/2}, \tag{29}$$

where we have used the Fourier representation of the Sobolev norm, $\int |h^{ft\prime}(t)|^2 dt \;=\; 2\pi\cdot \int |x\, h(x)|^2 dx$, and the Parseval identity. As both $\tilde g^{ft}$ and its (weak) derivative are supported on $[-\nu', \nu']$, the terms occurring in (29) have the upper bounds

$$\Big(\alpha_n^2 \cdot \int_{-\nu'}^{\nu'} \big|\varphi^{ft}(t/b_n)\big|^2 \big|Q_{K_n}^{ft}(t)\big|^2 dt\Big)^{1/2}, \qquad \Big(\alpha_n^2 b_n^{-2} \cdot \int_{-\nu'}^{\nu'} \big|\varphi^{ft\prime}(t/b_n)\big|^2 \big|Q_{K_n}^{ft}(t)\big|^2 dt\Big)^{1/2},$$
$$\Big(\alpha_n^2 \cdot \int_{-\nu'}^{\nu'} \big|\varphi^{ft}(t/b_n)\big|^2 \big|Q_{K_n}^{ft}{}'(t)\big|^2 dt\Big)^{1/2}. \tag{30}$$

As we have $\big|\varphi^{ft}(t/b_n)\big| \le 1$ and $\big|\varphi^{ft\prime}(t)\big| \le \int_{-1}^1 |x\varphi(x)| dx \le 1$, for all t, the terms in (30) are bounded above by

$$O(\alpha_n)\cdot \Big(\int_{-\nu'}^{\nu'} \big|Q_{K_n}^{ft}(t)\big|^2 dt \,+\, \int_{-\nu'}^{\nu'} \big|Q_{K_n}^{ft}{}'(t)\big|^2 dt\Big)^{1/2}. \tag{31}$$

We use the Taylor expansion of $Q_{K_n}^{ft}$ as in (3). Due to

$$\int_{-S_n/2}^{S_n/2} x^l \tilde P_n(2x/S_n) dx \;=\; 0\,,$$

for all non-negative integers $l < K_n$, the Taylor polynomial of $Q_{K_n}^{ft}$ with degree $K_n - 1$ vanishes. According to (4), we have

$$\big|Q_{K_n}^{ft}(t)\big| \;\le\; O(1)\cdot \big|S_n\, e\, t \,/\, K_n\big|^{K_n},$$

and, by considering the derivative of the residual term, we obtain

$$\big|Q_{K_n}^{ft}{}'(t)\big| \;\le\; O(1)\cdot \big|S_n\, e\, t \,/\, (K_n-1)\big|^{K_n-1}.$$

Therefore, (31) is bounded above by

$$O\big(\big|S_n\, e\, \nu' \,/\, (K_n-1)\big|^{(K_n-1)}\big) \;=\; O\Big(\exp\big\{(K_n-1)\cdot \big[\ln(S_n e \nu') - \ln(K_n-1)\big]\big\}\Big).$$

Finally we see that (28) is satisfied by the selection

$$K_n \;=\; 1 + c\cdot \ln n / \ln\ln n\,, \tag{32}$$

with $c > 1$ suitably large. Hence, the optimal rate of convergence is not faster than

$$\|f - f_n\|^2 \;=\; \alpha_n^2 \cdot \int \Big| \int b_n \varphi(b_n y) Q_{K_n}(x-y) dy\Big|^2 dx$$
$$\ge\; \text{const.}\cdot b_n^{-2\beta} K_n \cdot \int \Big| \int b_n' \varphi(b_n' y) \tilde P_{K_n}(x-y)\cdot \chi_{[-1,1]}(x-y) dy\Big|^2 dx\,, \tag{33}$$

when writing $b_n' = S_n b_n/2$ and inserting (27). We introduce some disjoint intervals $[s_j, t_j]$, $j = 1, \ldots, l_n$, which satisfy the properties given in Lemma 1. Therefore, we have $l_n > c_1 K_n$ and $\big|\tilde{P}_{K_n}(x)\big|^2 \geq c_2 K_n^{-1}$, $\forall x \in [s_j, t_j]$, $\forall j$. We set

$$b_n = 8K_n/(c_0 S_n)\,,$$

referred to c_0 in Lemma 1. Writing $s_j' = s_j + 1/b_n'$ and $t_j' = t_j - 1/b_n'$, we construct the interval $[s_j', t_j']$ having the length $\geq (c_0/2)K_n^{-1}$. As the sign of $\tilde{P}_{K_n}(x)$ does not change on $x \in [s_j, t_j]$, we may derive the following lower bound on (33),

$$\begin{aligned}
&\text{const.} \cdot b_n^{-2\beta} K_n \cdot \sum_{j=1}^{l_n} \int_{[s_j', t_j']} \Big| \int b_n' \varphi(b_n' y) \tilde{P}_{K_n}(x-y) \cdot \chi_{[-1,1]}(x-y) dy \Big|^2 dx \\
&\geq \text{const.} \cdot b_n^{-2\beta} K_n \cdot \sum_{j=1}^{l_n} (t_j' - s_j') \Big(\int b_n' \varphi(b_n' y) dy \Big)^2 \cdot \inf_{\xi \in [s_j' - 1/b_n', t_j' + 1/b_n']} \big|\tilde{P}_{K_n}(\xi)\big|^2 \\
&\geq \text{const.} \cdot b_n^{-2\beta} K_n \cdot \sum_{j=1}^{l_n} (c_0/2) K_n^{-1} \cdot \inf_{\xi \in [s_j, t_j]} \big|\tilde{P}_{K_n}(\xi)\big|^2 \\
&\geq \text{const.} \cdot b_n^{-2\beta} \cdot c_1 K_n \cdot c_2 K_n^{-1} \geq \text{const.} \cdot b_n^{-2\beta} \geq \text{const.} \cdot (K_n/S_n)^{-2\beta}\,.
\end{aligned}$$

Inserting (32) gives us the lower bound as stated in the theorem. ■

References

Bretagnolle, J. & Huber, C. (1979), Estimation des densités, risque minimax, *Z. Wahrsch. Verw. Gebiete*, **47**, 119–137.

Butucea, C. & Matias, C. (2005), Minimax estimation of the noise level and of the deconvolution density in a semiparametric convolution model, *Bernoulli*, **11**, 309–340.

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

Carroll, R.J., Ruppert, D. & Stefanski, L.A. (1995), *Measurement Error in Nonlinear Models.* Chapman and Hall, London.

Delaigle, A. & Gijbels, I. (2002), Estimation of integrated squared density derivatives from a contaminated sample, *J. R. Statist. Soc., B*, **64**, 869 – 886.

Delaigle, A. & Gijbels, I. (2004a), Bootstrap bandwidth selection in kernel density estimation from a contaminated sample, *Ann. Inst. Statist. Math.*, **56**, 19 – 47.

Delaigle, A. & Gijbels, I. (2004b), Comparison of data-driven bandwidth selection procedures in deconvolution kernel density estimation, *Comp. Statist. Data Anal.*, **45**, 249 – 267.

Devroye, L. (1989), Consistent deconvolution in density estimation, *Canad. J. Statist.*, **17**, 235–239.

Efromovich, S. (1997), Density estimation for the case of supersmooth measurement error, *J. Amer. Statist. Assoc.*, **92**, 526-535.

Fan, J. (1991a), On the optimal rates of convergence for non-parametric deconvolution problems, *Ann. Statist.*, **19**, 1257–1272.

Fan, J. (1991b), Global behavior of deconvolution kernel estimates, *Statist. Sinica*, **1**, 541–551.

Goldenshluger, A. (2002), Density deconvolution in the circular structural model, *J. Multivariate Anal.*, **84**, 350–375.

Fan, J. (1993), Adaptively local one-dimensional subproblems with application to a deconvolution problem, *Ann. Statist.*, **21**, 600–610.

Hall, P. & Qiu, P. (2005), Discrete-transform approach to deconvolution problems, *Biometrika*, **92**, 135–148.

Hall, P. & Meister, A. (2005), A ridge-parameter approach to deconvolution, *technical report*, Australian National University, Canberra, Australia.

Hesse, C.H. (1999), Data-driven deconvolution, *J. Nonparam. Statist.*, **10**, 343–373

Hesse, C.H. & Meister, A. (2004), Optimal iterative density deconvolution, *J. Nonparam. Statist.*, **16**, 879–900.

Koepf, W. (1998), *Hypergeometric Summation: An Algorithmic Approach to Summation and Special Function Identities*, Vieweg, Braunschweig, Germany.

Lukacs, E. (1970), *Characteristic Functions*, 2nd ed. Griffin, London.

Meister, A. (2004), On the effect of misspecifying the error density in a deconvolution problem, *Canad. J. Statist.*, **32**, 439–449.

Meister, A. (2006a), Density estimation with normal measurement error with unknown variance, to appear in *Statist. Sinica*.

Meister, A. (2006b), Non-estimability in spite of identifiability in density deconvolution, accepted for publication in *Math. Meth. Statist.*

Neumann, M.H. (1997), On the effect of estimating the error density in nonparametric deconvolution, *J. Nonparam. Statist.*, **7**, 307–330.

Pensky, M. & Vidakovic, B. (1999), Adaptive wavelet estimator for nonparametric density deconvolution, *Ann. Statist.*, **27**, 2033–2053.

Stefanski, L.A. & Carroll, R.J. (1990), Deconvoluting kernel density estimators, *Statistics*, **21**, 169–184.

Szegö, G. (1936), Inequalities for the zeros of Legendre polynomials and related functions, *Trans. Amer. Math. Soc.*, **39**, 1–17.

Alexander Meister
Pfaffenwaldring 57
70569 Stuttgart
Germany
E-Mail: `meistear@mathematik.uni-stuttgart.de`
WWW: `http://www.isa.uni-stuttgart.de/AbMathStat/Meister/index.html`

Erschienene Preprints ab Nummer 2004/001

Komplette Liste: `http://www.mathematik.uni-stuttgart.de/preprints`

2004/001 *Walk, H.:* Strong Laws of Large Numbers by Elementary Tauberian Arguments.

2004/002 *Hesse, C.H., Meister, A.:* Optimal Iterative Density Deconvolution: Upper and Lower Bounds.

2004/003 *Meister, A.:* On the effect of misspecifying the error density in a deconvolution problem.

2004/004 *Meister, A.:* Deconvolution Density Estimation with a Testing Procedure for the Error Distribution.

2004/005 *Efendiev, M.A., Wendland, W.L.:* On the degree of quasiruled Fredholm maps and nonlinear Riemann-Hilbert problems.

2004/006 *Dippon, J., Walk, H.:* An elementary analytical proof of Blackwell's renewal theorem.

2004/007 *Mielke, A., Zelik, S.:* Infinite-dimensional hyperbolic sets and spatio-temporal chaos in reaction-diffusion systems in $\mathbb{R}^n$.

2004/008 *Exner, P., Linde, H., Weidl T.:* Lieb-Thirring inequalities for geometrically induced bound states.

2004/009 *Ekholm, T., Kovarik, H.:* Stability of the magnetic Schrödinger operator in a waveguide.

2004/010 *Dillen, F., Kühnel, W.:* Total curvature of complete submanifolds of Euclidean space.

2004/011 *Afendikov, A.L., Mielke, A.:* Dynamical properties of spatially non-decaying 2D Navier-Stokes flows with Kolmogorov forcing in an infinite strip.

2004/012 *Pöschel, J.:* Hill's potentials in weighted Sobolev spaces and their spectral gaps.

2004/013 *Dippon, J., Walk, H.:* Simplified analytical proof of Blackwell's renewal theorem.

2004/014 *Kühnel, W.:* Tight embeddings of simply connected 4-manifolds.

2004/015 *Kühnel, W., Steller, M.:* On closed Weingarten surfaces.

2004/016 *Leitner, F.:* On pseudo-Hermitian Einstein spaces.

2004/017 *Förster, C., Östensson, J.:* Lieb-Thirring Inequalities for Higher Order Differential Operators.

2005/001 *Mielke, A.; Schmid, F.:* Vortex pinning in super-conductivity as a rate-independent model

2005/002 *Kimmerle, W.; Luca, F., Raggi-Cárdenas, A.G.:* Irreducible Components of the Burnside Ring

2005/003 *Höfert, C.; Kimmerle, W.:* On Torsion Units of Integral Group Rings of Groups of Small Order

2005/004 *Röhrl, N.:* A Least Squares Functional for Solving Inverse Sturm-Liouville Problems

2005/005 *Borisov, D.; Ekholm, T; Kovarik, H.:* Spectrum of the Magnetic Schrödinger Operator in a Waveguide with Combined Boundary Conditions

2005/006 *Zelik, S.:* Spatially nondecaying solutions of 2D Navier-Stokes equation in a strip

2005/007 *Meister, A.:* Deconvolving compactly supported densities

2005/008 *Förster, C., Weidl, T.:* Trapped modes for an elastic strip with perturbation of the material properties