Statistica Sinica **16**(2006), 195-211

# DENSITY ESTIMATION WITH NORMAL MEASUREMENT ERROR WITH UNKNOWN VARIANCE

Alexander Meister

*Universität Stuttgart*

*Abstract:* This paper deals with the problem of estimating a density based on observations which are contaminated by a normally distributed error whose variance is unknown. In the case of a completely unknown error variance, the impossibility of a uniformly consistent estimation is shown; however, a semi-uniformly consistent estimator is constructed under nonparametric smoothness conditions on the target density, and its rates are studied. If, in contrast, the error variance can be located in a known compact interval, we derive uniform consistency for this estimator which achieves nearly optimal rates. Simulations show the practical merit of the estimator.

*Key words and phrases:* Deconvolution, errors-in-variables, inversion problems, nonparametric estimation, reconstruction.

## 1. Introduction

The problem of reconstructing a density based on contaminated data has become a famous and widely studied topic in nonparametric density estimation. Several papers (e.g., Carroll and Hall (1988), Devroye (1989), Fan (1991), Fan (1993), Hesse (1999) and Liu and Taylor (1989)) are concerned with deconvolution estimation. In the basic problem, some empirical data $Y_1, \ldots, Y_n$ satisfying

$$Y_j = X_j + \varepsilon_j$$

are observed; the $\varepsilon_j$'s represent the error or the contamination, which possesses density $g$, and the $X_j$'s denote those random variables whose density $f$ is to be estimated. The components $X_1, \varepsilon_1, \ldots, X_n, \varepsilon_n$ are independent. Deterministic nonparametric knowledge about the densities $f$ and $g$, such as smoothness conditions, is usually given. In the current note, this knowledge is expressed by density classes $\mathcal{F}$ and $\mathcal{G}$ with $f \in \mathcal{F}$, $g \in \mathcal{G}$.

In the classical approach, the error density $g$ is assumed to be exactly known. Since this condition is not usually realistic, some papers dealing with an imperfectly known error density have been published. In the framework of Efromovich (1997) and Neumann (1997), the error density is unknown but it can be estimated based on additional direct observations which come from the error density. In

Meister (2004a), a testing procedure for two possible densities competing to be the error density is considered while $\mathcal{F}$ is assumed to be a special nonparametric smoothness class. Butucea and Matias (2004) introduce a uniformly consistent estimation procedure in a problem in which the error variance is unknown but restricted to a known compact interval.

In some situations, one might not even know such an interval. Hence, in the current note, we consider normal error densities with known mean and completely unknown variance. The new deconvolution problem has

$$\begin{aligned} \mathcal{F} &= \{f \text{ density } : C_2|t|^{-\beta} \geq |\psi_f(t)| \geq C_1|t|^{-\beta}, \text{ for all } t \text{ with } |t| \geq T > 0\}, \\ \mathcal{G} &= \{\mathrm{N}(\mu,\sigma^2) : \sigma^2 > 0\}, \end{aligned} \tag{1}$$

with constants $C_2 > C_1 > 0$, $\beta > 1$, $T > 0$ and $\mu$, which can be selected arbitrarily as long as $\mathcal{F}$ is not empty. Here $\mu$ is assumed to be known while the other parameters occurring in the definition of $\mathcal{F}$ might not be given in many applications, so we do not use them in the estimators' construction. Densities like those in $\mathcal{F}$ are called ordinary smooth in the notation of Fan (1991). As in Butucea and Matias (2004), a lower bound condition on the Fourier transform of the densities in $\mathcal{F}$ is needed to guarantee identifiability of the estimation problem, i.e., the unique reconstructability of the density $f$ from the observed density $h = f * g$. Ignorance of $\sigma^2$ necessitates more restrictive conditions on the target density, compared to the classical deconvolution problem. However $\mathcal{F}$, as defined in (1), contains a large class of densities in spite of the condition $\beta > 1$, which Butucea and Matias (2004) also assume. In order to underline the variety in $\mathcal{F}$ we give the following motivation. Assume a parameterized family of densities $\{f_\phi : \phi \in [\phi_0, \phi_1]\}$ with positive-valued Fourier transforms is included into $\mathcal{F}$. Then one can easily check that all densities $f$ with a randomly weighted parameter $\phi$, i.e., $f = \int f_\phi dQ(\phi)$ for any probability measure $Q$ on $[\phi_0, \phi_1]$, lie in $\mathcal{F}$. Hence $\mathcal{F}$ contains a nonparametric set of densities. As an example, consider the Laplace densities $f_\phi(x) = (0.5/\phi)\exp(-|x|/\phi)$ with Fourier transform $\psi_{f_\phi}(t) = 1/(1 + \phi^2 t^2)$; the $f_\phi$ are members of $\mathcal{F}$ for $\beta = 2$ and appropriate constants $C_1$, $C_2$ and $T$, if $\phi$ is restricted to some compact interval $[\phi_0, \phi_1]$. Furthermore, $f \in \mathcal{F}$ implies that all translations $f(\cdot - \mu)$ are in $\mathcal{F}$. Another aspect which emphasizes the size of $\mathcal{F}$ is the lower bound on the rates of convergence which we derive in Section 2.

We are mainly concerned with the estimation problem (1). However, we also study how well our estimators perform under the assumption of a given upper bound $\sigma_0^2 > 0$ for the error variance, leading to the estimation problem

$$\begin{aligned} \mathcal{F} &= \{f \text{ density } : C_2|t|^{-\beta} \geq |\psi_f(t)| \geq C_1|t|^{-\beta}, \text{ for all } t \text{ with } |t| \geq T > 0\} \\ \mathcal{G} &= \{\mathrm{N}(\mu,\sigma^2) : \sigma^2 \in (0, \sigma_0^2]\}. \end{aligned} \tag{2}$$

## 2. Consistent Estimation

First we describe the estimation method. We use the absolute empirical Fourier transform as an important tool, defined by

$$\hat{\varphi}_n(t) = \Big| n^{-1} \sum_{j=1}^{n} \exp(itY_j) \Big|. \tag{3}$$

In the sequel, $(k_n)_{n\in\mathbb{N}}$, $(\omega_n)_{n\in\mathbb{N}}$ and $(\sigma_n^2)_{n\in\mathbb{N}}$ denote sequences of positive numbers which will be determined later. Due to the upper and lower bound conditions in $\mathcal{F}$, we are able to estimate the error variance consistently. Therefore, setting $\tilde{\sigma_n^2} = -2\, k_n^{-2} \ln(\hat{\varphi}_n(k_n)/(C_1' k_n^{-\beta'}))$ with constants $\beta' > 1$, $C_1' > 0$, we can derive an explicit estimator for the error variance by truncating $\tilde{\sigma_n^2}$ to the interval $[0, \sigma_n^2]$:

$$\hat{\sigma_n^2} = \begin{cases} 0, & \text{if } \tilde{\sigma_n^2} < 0, \\ \tilde{\sigma_n^2}, & \text{if } \tilde{\sigma_n^2} \in [0, \sigma_n^2], \\ \sigma_n^2, & \text{if } \tilde{\sigma_n^2} > \sigma_n^2. \end{cases} \tag{4}$$

Note that $\beta' > 1$ and $C_1' > 0$ may be arbitrary. As motivated by their notation, they should correspond to $\beta$ and $C_1$ which, however, are not stipulated to be known and, therefore, might be misspecified. We use

*Selection rule* (S1): If we know $C_1$ and $\beta$ we choose $\beta' = \beta$, $C_1' = C_1$ and $k_n = \omega_n$; otherwise, we set $\omega_n = k_n / \ln k_n$. In any case, we select $k_n \to \infty$.

We construct our density estimator by replacing the error variance by the empirical variance (4) for the estimation problems (1) and (2):

$$\hat{f}_n(x) = (2\pi)^{-1} \int_{-\omega_n}^{\omega_n} \exp\Big( -it(x+\mu) + \frac{1}{2} \hat{\sigma_n^2} t^2 \Big) n^{-1} \sum_{j=1}^{n} \exp(itY_j)\, dt. \tag{5}$$

Now our investigation concentrates on the asymptotic properties of (5), related to its MISE (=mean integrated squared error).

**Lemma 1.** *In* $\mathcal{G}_n = \{N(\mu, \sigma^2) : \sigma^2 \in (0, \sigma_n^2]\}$, *the MISE of* (5) *satisfies*

$$\sup_{g\in\mathcal{G}_n} \sup_{f\in\mathcal{F}} \mathrm{E}_{f,g} \|\hat{f}_n - f\|^2_{L_2(\mathbb{R})} \le B + V + E,$$

*where, with* $d_n := 2 \ln\big((2C_2/C_1)\big)\omega_n^{-2}$,

$$B \le \textit{const.}\, \omega_n^{1-2\beta}, \qquad V \le \textit{const.}\, n^{-1} \omega_n \exp(\sigma_n^2 \omega_n^2),$$

$$E \le \textit{const.} \sup_{f\in\mathcal{F}} \sup_{g\in\mathcal{G}} \Big( \omega_n \int_{-1}^{1} |\psi_f(\omega_n s)|^2 s^4 ds + \omega_n \exp(\sigma_n^2 \omega_n^2) \int_{-1}^{1} |\psi_f(\omega_n s)|^2 \times P_{f,g}(|\hat{\sigma_n^2} - \sigma^2| > d_n) ds \Big),$$

**Lemma 2.** *Let $d_n$ and $\mathcal{G}_n$ be as in Lemma* 1. *Under* (S1),

$$\sup_{f\in\mathcal{F}}\sup_{g\in\mathcal{G}_n} P_{f,g}\big(|\hat{\sigma_n^2}-\sigma^2|>d_n\big)\le const.k_n^{2\beta}\exp\frac{(\sigma_n^2k_n^2)}{n}.$$

There are also negative results. They involve the Sobolev space $W^1 := \{f \in L_2(\mathbb{R}) : f' \in L_2(\mathbb{R})\}$.

**Lemma 3.** *Consider an arbitrary deconvolution problem with density classes $\mathcal{F}\subseteq L_2(\mathbb{R})$ and $\mathcal{G}$. Assume that the $L_2(\mathbb{R})$-norm of the densities in $\mathcal{F}$ possesses a uniform upper bound $C$. Let $(f_n)_n, (\tilde{f}_n)_n \subseteq \mathcal{F}$ and $(g_n)_n, (\tilde{g}_n)_n \subseteq \mathcal{G}$ be density sequences and define $(h_n)_n = (f_n * g_n)_n$ and $(\tilde{h}_n)_n = (\tilde{f}_n * \tilde{g}_n)_n$. Assume that $\psi_{h_n}$ and $\psi_{\tilde{h}_n}$ are in $W^1$. If*

$$\Big(nR_n^{\frac{3}{4}}\|h_n-\tilde{h}_n\|_{L_2(\mathbb{R})}+nR_n^{-\frac{1}{4}}(\|\psi'_{h_n}\|_{L_2(\mathbb{R})}+\|\psi'_{\tilde{h}_n}\|_{L_2(\mathbb{R})})\Big)\|f_n-\tilde{f}_n\|^{-2}_{L_2(\mathbb{R})}\underset{n\to\infty}{\longrightarrow}0 \tag{6}$$

*for an appropriate sequence $(R_n)_n$, then the MISE of an arbitrary estimator $\hat{f}_n$ based on $Y_1,\ldots,Y_n$ satisfies*

$$\sup_{g\in\mathcal{G}}\sup_{f\in\mathcal{F}}\mathrm{E}_{f,g}\|\hat{f}_n(Y_1,\ldots,Y_n)-f\|^2_{L_2(\mathbb{R})}\ge const.\,\|f_n-\tilde{f}_n\|^2_{L_2(\mathbb{R})}.$$

Lemmas 1−3 provide the necessary tools for studying our estimation problems. We would like the estimator to be uniformly consistent for (1), but have the following instead.

**Theorem 1.** *Assume $T^\beta \ge C_2(\beta+1)$ holds in* (1). *Then there is no estimator $\hat{f}_n$ with $\sup_{g\in\mathcal{G}}\sup_{f\in\mathcal{F}}E_{f,g}\|\hat{f}_n(Y_1,\ldots,Y_n)-f\|^2_{L_2(\mathbb{R})}\underset{n\to\infty}{\longrightarrow}0$.*

Theorem 1 says that uniformly consistent estimation is impossible. This may look surprising as the density $f$ can be identified in problem (1). However identifiability, in general, does not imply the existence of a uniformly consistent estimator, not even the existence of a consistent estimator, as was shown in Meister (2003).

Nevertheless, (5) satisfies a weaker version of consistency; we call it semi-uniform consistency, i.e., uniform related to $\mathcal{F}$ but individual related to $\mathcal{G}$.

**Theorem 2.** *Consider* (1) *and take* (5) *with $(\sigma_n^2)_{n\in\mathbb{N}} = (0.25(\ln\ln n))_{n\in\mathbb{N}}$, $(k_n)_{n\in\mathbb{N}} = ((\ln n/\ln\ln n)^{1/2})_n$ and $(\omega_n)_n$ according to selection rule* (S1). *Then, for any $g\in\mathcal{G}$, $\sup_{f\in\mathcal{F}}E_{f,g}\|\hat{f}_n-f\|^2_{L_2(\mathbb{R})}$ is bounded above by*

| | $\beta < 2.5$ | $\beta > 2.5$ | $\beta = 2.5$ |
|---|---|---|---|
| (i) | $(\ln\ln n)^{\beta-0.5}(\ln n)^{0.5-\beta}$ | $(\ln\ln n)^2(\ln n)^{-2}$ | $(\ln\ln n)^3(\ln n)^{-2}$ |
| (ii) | $(\ln\ln n)^{3\beta-1.5}(\ln n)^{0.5-\beta}$ | $(\ln\ln n)^4(\ln n)^{-2}$ | $(\ln\ln n)^5(\ln n)^{-2}$, |

*multiplied by a constant depending on $g$;* (i) *is the case of known $\beta$ and $C_1$, else* (ii).

Now consider the estimator (5) in the problem (2). Under slightly different smoothness conditions on $\mathcal{F}$, Butucea and Matias (2004) derive a consistent estimator. However, the selection of the bandwidth sequence corresponding to $(\omega_n)_n$ in their Theorem 6 is non-adaptive, i.e., the sequence depends on a parameter $\beta'$ of $\mathcal{F}$. Here we show that our procedure allows a bandwidth selection that is independent of the smoothness parameters of $\mathcal{F}$ as usually required in deconvolution estimation. As we focus on normal measurement error, no data-driven procedure is needed for the selection of the bandwidth sequence. Concerning practical issues of the bandwidth choice and kernel deconvolution estimation in general, we mention the papers of Delaigle and Gijbels (2004a, 2004b).

**Theorem 3.** *Consider* (2)*. Select $(\sigma_n^2)_n = (\sigma_0^2)_{n\in\mathbb{N}}$, $(k_n)_{n\in\mathbb{N}} = (u(\ln n)^{0.5})_n$ with $u = 1/(2\sigma_0)$, and $(\omega_n)_n$ according to* (S1)*. Then,* $\sup_{g\in\mathcal{G}}\sup_{f\in\mathcal{F}} E_{f,g}\|\hat{f}_n - f\|^2_{L_2(\mathbb{R})}$ *is bounded above by*

| | $\beta < 2.5$ | $\beta > 2.5$ | $\beta = 2.5$ |
|---|---|---|---|
| (i) | $(\ln n)^{0.5-\beta}$ | $(\ln n)^{-2}$ | $(\ln\ln n)(\ln n)^{-2}$ |
| (ii) | $(\ln\ln n)^{2\beta-1}(\ln n)^{0.5-\beta}$ | $(\ln\ln n)^{4}(\ln n)^{-2}$ | $(\ln\ln n)^{5}(\ln n)^{-2}$, |

*multiplied by a constant. The cases* (i) *and* (ii) *are as in Theorem* 2.

We see that rates are optimal or nearly optimal up to multiplication of an iterated logarithmic term in finite power. Since the smoothness classes of Butucea and Matias (2004) are not subsets of our class $\mathcal{F}$, we cannot use those results.

**Theorem 4.** *Let $(T/e)^\beta \geq C_2$ in* (2)*. The MISE of an arbitrary estimator $\hat{f}_n$ based on $Y_1,\ldots,Y_n$ satisfies*

$$\sup_{g\in\mathcal{G}}\sup_{f\in\mathcal{F}} \mathrm{E}_{f,g}\|\hat{f}_n(Y_1,\ldots,Y_n) - f\|^2_{L_2(\mathbb{R})} \geq const. \begin{cases} (\ln n)^{0.5-\beta}, & \text{if } \beta < 2.5, \\ (\ln n)^{-2}, & \text{if } \beta \geq 2.5. \end{cases}$$

For $\beta < 2.5$, the rates in Theorem 4 correspond to those derived by Fan (1993) for Hölder classes of densities under the assumption of a known error density. When $\beta > 2.5$, the rates deteriorate to Fan's results; we pay the price for the unknown error variance. These rates also correspond to those derived by Butucea and Matias (2004) for their smoothness classes.

Note that misspecification of the smoothness parameter $\beta$ is much less sensitive than misspecification of the error variance in the classical deconvolution problem. While inserting $\beta$ and $C_1$ erroneously into (5) causes a slight deterioration of the rates but keeps consistency, any misspecification of the error variance leads to inconsistency and, in some circumstances, even to the divergence of the MISE to infinity (see Meister (2004b)). Hence, compared to the classical deconvolution procedure, (5) shows greater robustness properties.

## 3. Proofs

**Proof of Lemma 1.** There is a $N$ so that $\omega_n > T$ holds for all $n \geq N$. Hence the upper and lower bound of the Fourier transform can be used. We start by deriving an upper bound using Parseval's identity and Fubini's theorem.

$$\begin{aligned}
&\sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}} \mathrm{E}_{f,g}\|\hat{f}_n - f\|^2_{L_2(\mathbb{R})} \\
&= (2\pi)^{-1} \sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}} \Big( \int_{-\omega_n}^{\omega_n} \mathrm{E}_{f,g}\Big|\frac{1}{n}\sum_{j=1}^n \exp\big(it(Y_j-\mu)\big)\exp(\frac{1}{2}\hat{\sigma_n^2}t^2) \\
&\quad -\psi_f(t)\Big|^2 dt + \int_{|t|\geq\omega_n} |\psi_f(t)|^2 dt\Big) \\
&\leq (2\pi)^{-1}\Big( \sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}} 2\int_{\omega_n}^{\infty} |\psi_f(t)|^2 dt \\
&\quad + \sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}} \int_{-\omega_n}^{\omega_n} 2\mathrm{E}_{f,g}\Big|\exp(-i\mu t + \frac{1}{2}\hat{\sigma_n^2}t^2)\Big(n^{-1}\sum_{j=1}^n \exp(itY_j) - \psi_h(t)\Big)\Big|^2 dt \\
&\quad + \sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}} \int_{-\omega_n}^{\omega_n} 2\mathrm{E}_{f,g}\Big|\psi_h(t)/\big(\exp(it\mu - \frac{1}{2}\hat{\sigma_n^2}t^2)\big) - \psi_f(t)\Big|^2 dt\Big).
\end{aligned}$$

We show that the first addend is bounded above by $B$, the second addend by $V$ and the third one by $E$, as stated in the theorem. The bound on the bias term $B$ can be determined in the usual way by using the upper bound of the Fourier transform from $\mathcal{F}$.

The variance term $V$ can be bounded above via (4):

$$V \leq \sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}} \int_{-\omega_n}^{\omega_n} 2\exp(\sigma_n^2 t^2)\mathrm{E}_{f,g}\Big|\frac{1}{n}\sum_{j=1}^n \exp(itY_j) - \psi_h(t)\Big|^2 dt \leq \frac{4}{n}\omega_n \exp(\sigma_n^2\omega_n^2).$$

Term $E$ does not occur in classical deconvolution estimation and, hence, it has to be studied precisely. Substituting $s = \omega_n^{-1}t$, $E$ is

$$\begin{aligned}
&\sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}} \int_{-\omega_n}^{\omega_n} 2\mathrm{E}_{f,g}\Big|\psi_f(t)/\big(\exp\big(i\mu t - i\mu t - \frac{1}{2}(\hat{\sigma_n^2} - \sigma^2)t^2\big)\big) - \psi_f(t)\Big|^2 dt \\
&\leq 2\sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}} \omega_n \int_{-1}^{1} \mathrm{E}_{f,g}\big|\exp\Big(\frac{1}{2}|\hat{\sigma_n^2} - \sigma^2|\omega_n^2 s^2\Big) - 1\big|^2 |\psi_f(\omega_n s)|^2 ds. \qquad (7)
\end{aligned}$$

Let $s \in [-1,1]$ and $f \in \mathcal{F}$, $g \in \mathcal{G}_n$. Applying $(d_n)_{n\in\mathbb{N}}$, we have

$$
\begin{aligned}
&\mathrm{E}_{f,g}\big|\exp\big(\frac{1}{2}|\hat{\sigma_n^2}-\sigma^2|\omega_n^2 s^2\big)-1\big|^2\\
&=\mathrm{E}_{f,g}\Big(\big|\exp\big(\frac{1}{2}|\hat{\sigma_n^2}-\sigma^2|\omega_n^2 s^2\big)-1\big|^2\chi_{(|\hat{\sigma_n^2}-\sigma^2|\le d_n)}\Big)\\
&\quad+\mathrm{E}_{f,g}\Big(\big|\exp\big(\frac{1}{2}|\hat{\sigma_n^2}-\sigma^2|\omega_n^2 s^2\big)-1\big|^2\chi_{(|\hat{\sigma_n^2}-\sigma^2|> d_n)}\Big)\\
&\le\big|\exp\big(\frac{1}{2}d_n\omega_n^2 s^2\big)-1\big|^2+\exp(\sigma_n^2\omega_n^2)\mathrm{P}_{f,g}(|\hat{\sigma_n^2}-\sigma^2|>d_n).
\end{aligned}
$$

Since $s^2 \le 1$, the $(d_n\omega_n^2 s^2)_{n\in\mathbb{N}}$ is bounded by $2\ln(2C_2/C_1)$ independent of $s$. Now $g(x) = (\exp(x)-1)/x$ is continuous on $[0, \ln(2C_2/C_1)]$ if one defines $g(0) = 1$. So $g(x)$ is bounded above for $x \in [0, \ln(2C_2/C_1)]$. Hence, the sequence $(\big|(\exp(0.5d_n\omega_n^2 s^2)-1)/(0.5d_n\omega_n^2 s^2)\big|^2)_n$ also has an upper bound that is independent of $s$. The inequality sequence continues with

$$
\begin{aligned}
&|\exp\big(\frac{1}{2}d_n\omega_n^2 s^2\big)-1|^2+\exp(\sigma_n^2\omega_n^2)\mathrm{P}_{f,g}(|\hat{\sigma_n^2}-\sigma^2|>d_n)\\
&\le const.\, s^4+\exp(\sigma_n^2\omega_n^2)\mathrm{P}_{f,g}(|\hat{\sigma_n^2}-\sigma^2|>d_n). \qquad (8)
\end{aligned}
$$

Inserting (8) into (7) gives the bound stated in the lemma.

**Proof of Lemma 2.** The term $\sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}}\mathrm{P}_{f,g}(|\hat{\sigma_n^2}-\sigma^2|\ge d_n)$ is bounded above by two addends; we derive an upper bound for each of them. Focus on the first addend. Since the supremum is considered for $g \in \mathcal{G}_n$, we have $\sigma \in (0, \sigma_n]$. Setting $h = f * g$, we study

$$
\sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}}\mathrm{P}_{f,g}(\hat{\sigma_n^2}-\sigma^2\ge d_n)\le\sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}}\mathrm{P}_{f,g}\big(\hat{\varphi}_n(k_n)\le\alpha_n|\psi_h(k_n)|\big) \qquad (9)
$$

with $\alpha_n := (C_1'/C_1)k_n^{\beta-\beta'}\exp(-0.5k_n^2 d_n)$; we have used $|\psi_h(k_n)| \ge C_1 k_n^{-\beta}\exp(-0.5k_n^2\sigma^2)$ for that purpose. In the case of known parameters $\beta = \beta'$, $C_1 = C_1'$, $\alpha_n$ is $C_1/(2C_2) < 1$. Otherwise, respecting the parameter selection stated in the lemma, we have

$$
k_n^{|\beta-\beta'|}\exp(-0.5k_n^2 d_n)\ \longrightarrow_{n\to\infty}\ 0 \qquad (10)
$$

and, hence, $\alpha_n \to 0$. In both cases, the existence of a constant $c \in (0,1)$ is guaranteed so that (9) is bounded above by

$$
\begin{aligned}
&\sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}}\mathrm{P}_{f,g}\big(\hat{\varphi}_n(k_n)\le c|\psi_h(k_n)|\big)\\
&\le(1-c)^{-2}\sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}}|\psi_h(k_n)|^{-2}\mathrm{E}_h\Big|\frac{1}{n}\sum_{j=1}^{n}\exp(ik_nY_j)-\psi_h(k_n)\Big|^2\\
&\le const. k_n^{2\beta}\exp(\sigma_n^2k_n^2)n^{-1}.
\end{aligned}
$$

The second addend can be bounded in a similar way.

$$\sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}} \mathrm{P}_{f,g}(\hat{\sigma_n^2}-\sigma^2\le -d_n) \le \sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}} \mathrm{P}_{f,g}\big(\hat{\varphi_n}\ge\gamma_n|\psi_h(k_n)|\big) \tag{11}$$

with $\gamma_n := (C_1'/C_2)k_n^{\beta-\beta'}\exp(0.5k_n^2 d_n)$. In case $\beta'=\beta$ and $C_1'=C_1$, we have $\gamma_n=2$; otherwise, (10) implies $\gamma_n \longrightarrow_{n\to\infty} \infty$. Hence, in both cases, the existence of a constant $C>1$ may be assumed so that in (11),

$$\begin{aligned}&\sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}} P_{f,g}\big(\hat{\varphi_n}\ge C|\psi_h(k_n)|\big)\\ &\le (C-1)^2\sup_{g\in\mathcal{G}_n}\sup_{f\in\mathcal{F}}|\psi_h(k_n)|^{-2}\mathrm{E}_h\Big|\frac{1}{n}\sum_{j=1}^n \exp(ik_nY_j)-\psi_h(k_n)\Big|^2.\end{aligned}$$

This leads to the same upper bound as derived for the first addend.

**Proof of Lemma 3.** In view of the uniform upper bound $C$ of the $L_2(\mathbb{R})$-norm of the densities in $\mathcal{F}$, the $L_2(\mathbb{R})$-norm of the estimator $\hat{f}_n$ can be assumed to be uniformly bounded above by $C$ without loss of generality, since the $L_2(\mathbb{R})$-distance between the norm-truncated estimator

$$\tilde{f}_n = \begin{cases}\hat{f}_n, & \text{if } \|\hat{f}_n\|_{L_2(\mathbb{R})}\le C,\\ \frac{C}{\left(\|\hat{f}_n\|_{L_2(\mathbb{R})}\right)\hat{f}_n}, & \text{otherwise},\end{cases}$$

and the target density $f$ is not larger than the distance of $f$ and $\hat{f}_n$ almost surely. Here $C$ may be assumed to be known since we are considering the lower bound result. Now

$$\begin{aligned}&\sup_{g\in\mathcal{G}}\sup_{f\in\mathcal{F}} \mathrm{E}_{f,g}\|\hat{f}_n(Y_1,\ldots,Y_n)-f\|^2_{L_2(\mathbb{R})}\\
&\ge 0.5\left(\mathrm{E}_{\tilde{f}_n,\tilde{g}_n}\|\hat{f}_n(Y_1,\ldots,Y_n)-\tilde{f}_n\|^2_{L_2(\mathbb{R})}+\mathrm{E}_{f_n,g_n}\|\hat{f}_n(Y_1,\ldots,Y_n)-f_n\|^2_{L_2(\mathbb{R})}\right)\\
&=0.5\Big(\mathrm{E}_{\tilde{f}_n,\tilde{g}_n}\|\hat{f}_n(Y_1,\ldots,Y_n)-\tilde{f}_n\|^2_{L_2(\mathbb{R})}+\mathrm{E}_{\tilde{f}_n,\tilde{g}_n}\|\hat{f}_n(Y_1,\ldots,Y_n)-f_n\|^2_{L_2(\mathbb{R})}\\
&\quad -\mathrm{E}_{\tilde{f}_n,\tilde{g}_n}\|\hat{f}_n(Y_1,\ldots,Y_n)-f_n\|^2_{L_2(\mathbb{R})}+\mathrm{E}_{f_n,g_n}\|\hat{f}_n(Y_1,\ldots,Y_n)-f_n\|^2_{L_2(\mathbb{R})}\Big)\\
&\ge 0.5\Big(\mathrm{E}_{\tilde{f}_n,\tilde{g}_n}(\|\hat{f}_n(Y_1,\ldots,Y_n)-\tilde{f}_n\|^2_{L_2(\mathbb{R})}+\|\hat{f}_n(Y_1,\ldots,Y_n)-f_n\|^2_{L_2(\mathbb{R})})\\
&\quad -\big|-\mathrm{E}_{\tilde{f}_n,\tilde{g}_n}\|\hat{f}_n(Y_1,\ldots,Y_n)-f_n\|^2_{L_2(\mathbb{R})}+\mathrm{E}_{f_n,g_n}\|\hat{f}_n(Y_1,\ldots,Y_n)-f_n\|^2_{L_2(\mathbb{R})}\big|\Big)\end{aligned}$$

$$
\begin{aligned}
&\geq 0.5\Big(0.5\mathrm{E}_{\tilde{f}_n,\tilde{g}_n}\|\hat{f}_n(Y_1,\ldots,Y_n)-\tilde{f}_n-\hat{f}_n(Y_1,\ldots,Y_n)+f_n\|^2_{L_2(\mathbb{R})}\\
&\quad-\int\cdots\int\|\hat{f}_n(y_1,\ldots,y_n)-f_n\|^2_{L_2(\mathbb{R})}|h_n(y_1)\cdots h_n(y_n)-\tilde{h}_n(y_1)\cdots\tilde{h}_n(y_n)|\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad dy_1\cdots dy_n\Big)\\
&\geq 0.5\big(0.5\|\tilde{f}_n-f_n\|^2_{L_2(\mathbb{R})}-2C^2n\int|h_n(y)-\tilde{h}_n(y)|dy\big).
\end{aligned}
$$

We define the density

$$
\xi(t)=\begin{cases} c, & \text{if } |t|\leq 1,\\ c|t|^{-\frac{3}{2}}, & \text{if } |t|>1,\end{cases}
$$

with an appropriate constant $c$. Then, utilizing the Cauchy-Schwarz-inequality, the inequality sequence continues with

$$
\begin{aligned}
&\frac{1}{2}\big(\frac{1}{2}\|\tilde{f}_n-f_n\|^2_{L_2(\mathbb{R})}-2C^2n\int|h_n(y)-\tilde{h}_n(y)|dy\big)\\
&=\frac{1}{2}\big(\frac{1}{2}\|\tilde{f}_n-f_n\|^2_{L_2(\mathbb{R})}-2C^2n\int\frac{\sqrt{\xi(y)}|h_n(y)-\tilde{h}_n(y)|}{\sqrt{\xi(y)}}dy\big)\\
&\geq\frac{1}{2}\Big(\frac{1}{2}\|\tilde{f}_n-f_n\|^2_{L_2(\mathbb{R})}-2C^2n\Big(\int\frac{|h_n(y)-\tilde{h}_n(y)|^2}{\xi(y)}dy\Big)^{0.5}\Big).
\end{aligned}
$$

We have used the fact that $\xi$ integrates to 1. Now our goal is selecting the functions $f_n$, $\tilde{f}_n$ and $\tilde{g}_n$, $g_n$ so that

$$
n\Big(\int\frac{|h_n(y)-\tilde{h}_n(y)|^2}{\xi(y)}dy\Big)^{0.5}\|\tilde{f}_n-f_n\|^{-2}_{L_2(\mathbb{R})}\longrightarrow_{n\to\infty} 0. \tag{12}
$$

Then the MISE is greater or equal to $(1/8)\|\tilde{f}_n-f_n\|^2_{L_2(\mathbb{R})}$ for $n\geq N$ and for $N$ sufficiently large. Further calculation leads to

$$
\begin{aligned}
&n\big(\int\frac{|h_n(y)-\tilde{h}_n(y)|^2}{\xi(y)}dy\big)^{0.5}\\
&=n\Big(\int_{|y|\leq R_n}\frac{|h_n(y)-\tilde{h}_n(y)|^2}{\xi(y)}dy+\int_{|y|>R_n}\frac{|h_n(y)-\tilde{h}_n(y)|^2}{\xi(y)}dy\Big)^{0.5}\\
&\leq const.nR_n^{3/4}\|h_n-\tilde{h}_n\|_{L_2(\mathbb{R})}+2n\Big(\int_{|y|>R_n}\frac{|h_n(y)|^2}{\xi(y)}dy+\int_{|y|>R_n}\frac{|\tilde{h}_n(y)|^2}{\xi(y)}dy\Big)^{0.5},
\end{aligned}
$$

for any sequence $(R_n)_n$ tending to infinity.

Applying the Fourier-analytic results $\|\psi_{h'}\|_{L_2(\mathbb{R})} = \|\bullet\psi_h(\bullet)\|_{L_2(\mathbb{R})}$ and $\psi_{\psi_h} = 2\pi\, h(-\bullet)$ in $W^1$, we get

$$\begin{aligned} +\infty > \|\psi'_{h_n}\|^2_{L_2(\mathbb{R})} + \|\psi'_{\tilde{h}_n}\|^2_{L_2(\mathbb{R})} &= (2\pi)^{-1}\, \|\psi_{\psi'_{h_n}}\|^2_{L_2(\mathbb{R})} + (2\pi)^{-1}\|\psi_{\psi'_{\tilde{h}_n}}\|^2_{L_2(\mathbb{R})} \\ &= \int |t|^2 h_n(-t)^2 dt + \int |t|^2 \tilde{h}_n(-t)^2 dt = \int t^2 h_n(t)^2 dt + \int t^2 \tilde{h}_n(t)^2 dt. \quad (13) \end{aligned}$$

That condition of integrability is used in order to derive an upper bound for the terms above. We have

$$\begin{aligned} &\int_{|y|>R_n} \frac{|h_n(y)|^2}{\xi(y)} dy + \int_{|y|>R_n} \frac{|\tilde{h}_n(y)|^2}{\xi(y)} dy \\ &\leq const.\Big(\int_{|y|>R_n} |h_n(y)|^2 y^2 |y|^{-0.5} dy + \int_{|y|>R_n} |\tilde{h}_n(y)|^2 y^2 |y|^{-0.5} dy\Big) \\ &\leq const.\Big(R_n^{-0.5}\int_{|y|>R_n} |h_n(y)|^2 y^2 dy \;+\; R_n^{-0.5}\int_{|y|>R_n} |\tilde{h}_n(y)|^2 y^2 dy\Big) \\ &\leq const.R_n^{-0.5}\big(\|\psi'_{h_n}\|^2_{L_2(\mathbb{R})} + \|\psi'_{\tilde{h}_n}\|^2_{L_2(\mathbb{R})}\big). \end{aligned}$$

Hence, (12) holds if (6) is satisfied.

**Proof of Theorem 1.** Define the function

$$\varphi(t) = \begin{cases} \frac{C_2 T^{-\beta}-1}{T}|t| + 1, & \text{if } |t| \leq T, \\ C_2|t|^{-\beta}, & \text{otherwise.} \end{cases}$$

Notice that $\alpha(t) = 1 + \big((C_2T^{-\beta}-1)/T\big)t - C_1 t^{-\beta} = 0$ has at least one solution in $[T,+\infty)$ since $\alpha$ is continuous, $\alpha(T) = (C_2 - C_1)T^{-\beta} > 0$ and $\alpha(t) \longrightarrow_{t\to+\infty} -\infty$ hold. As the set of all solutions in $[T,+\infty)$ is closed, the minimum of the solutions in $[T,+\infty)$ exists $-$ let us call it $T^*$. Then we can define

$$\varphi^*(t) = \begin{cases} \frac{C_2 T^{-\beta}-1}{T}|t| + 1, & \text{if } |t| \leq T^*, \\ C_1|t|^{-\beta}, & \text{otherwise.} \end{cases}$$

Checking the conditions of Polya's criterion (see Durrett (1996, p.104)) and using $T^\beta \geq C_2(\beta+1)$, one can see that there are densities $f$ and $f^*$ so that $\varphi = \psi_f$ and $\varphi^* = \psi_{f^*}$. Notice that $\varphi$ and $\varphi^*$ agree on their restriction to $[-T,+T]$. Since the defining condition of $\mathcal{F}$ is also satisfied, we have $f, f^* \in \mathcal{F}$ with $f \neq f^*$. With respect to Lemma 3, one sets $f_n = f$, $\tilde{f}_n = f^*$ and $g_n = \tilde{g}_n = \mathrm{N}(\mu, n)$. Then, $\sup_{g\in\mathcal{G}}\sup_{f\in\mathcal{F}} \mathrm{E}_{f,g}\|\hat{f}_n(Y_1,\ldots,Y_n) - f\|^2_{L_2(\mathbb{R})} \geq const.\|f - f^*\|^2_{L_2(\mathbb{R})} > 0$ if

$$nR_n^{\frac{3}{4}}\|h_n - \tilde{h}_n\|_{L_2(\mathbb{R})} + nR_n^{-\frac{1}{4}}\big(\|\psi'_{h_n}\|_{L_2(\mathbb{R})} + \|\psi'_{\tilde{h}_n}\|_{L_2(\mathbb{R})}\big) \longrightarrow_{n\to\infty} 0 \qquad (14)$$

holds. Notice that $\psi_f$ is differentiable on $\mathbb{R}\backslash\{-T,0,T\}$, its derivative is square integrable and, hence, $\psi_f \in W^1$ is valid. We have

$$\|\psi'_{h_n}\|^2_{L_2(\mathbb{R})} \leq const.n^2 \int t^2 \exp(-nt^2)dt \leq const.\sqrt{n}.$$

Accordingly, we see that $\psi_{f^*}$ lies in $W^1$. So we get $\|\psi'_{\tilde{h}_n}\|^2_{L_2(\mathbb{R})} \leq const.\sqrt{n}$. Furthermore we have

$$\|\tilde{h}_n - h_n\|^2_{L_2(\mathbb{R})} \leq \pi^{-1} \int_T^\infty |\varphi(t) - \varphi^*(t)|^2 \exp(-nt^2)dt \leq O\big(\exp(-nT^2)\big).$$

Set $R_n = n^6$, then (14) is valid.

**Proof of Theorem 2.** First, fix $g \in \mathcal{G}$. The variance of $g$ is called $\sigma^2$. As $(\sigma_n)_{n\in\mathbb{N}} \to +\infty$ we can assume $n$ to be large enough so that $\sigma_n^2 > \sigma^2$. Hence, $g$ is a member of the sets $\mathcal{G}_n$ if $n > N$ for some $N \in \mathbb{N}$ sufficiently large. This implies $\sup_{f\in\mathcal{F}} \mathrm{E}_{f,g}\|\hat{f}_n - f\|^2_{L_2(\mathbb{R})} \leq \sup_{g\in\mathcal{G}_n} \sup_{f\in\mathcal{F}} \mathrm{E}_{f,g}\|\hat{f}_n - f\|^2_{L_2(\mathbb{R})}$. Therefore, Lemmas 1 and 2 can be used for the further calculation. The terms $B$ and $V$ as well as the second addend of term $E$, combined with Lemma 2, give the upper bound

$$\max\big\{\omega_n^{1-2\beta}\,,\, k_n^{2\beta+1}\exp\frac{(2\sigma_n^2 k_n^2)}{n}\big\}, \tag{15}$$

due to $k_n \geq \omega_n$, which follows from (S1). To derive an upper bound for the first addend of $E$, we have to distinguish between three cases: in the case of $\beta < 5/2$, we apply the upper bound of $|\psi_f|$ to $f \in \mathcal{F}$ to get

$$\omega_n \int_{-1}^1 |\psi_f(\omega_n s)|^2 s^4 ds \leq \omega_n^{1-2\beta} \int_{-1}^1 s^{4-2\beta} ds.$$

Hence, in this case, (15) is the upper bound of the MISE. By inserting the parameter sequences as in the lemma, we get the corresponding rates whether $\beta$ and $C_1$ are known or not.

If $\beta > 5/2$, we see that the first addend of $E$ has the bound

$$\omega_n \int_{-1}^1 |\psi_f(\omega_n s)|^2 s^4 ds \leq \omega_n^{-4} \int_{-\omega_n}^{\omega_n} t^4 |\psi_f(t)|^2 dt,$$

while the integral possesses an upper bound depending on neither $n$ nor $f$, due to the case condition. Again the upper bound of $|\psi_f|$, which is guaranteed by the membership of $f$ in $\mathcal{F}$, leads to the upper bound $\max\big\{\omega_n^{-4}\,,\,(\ln n)^{\beta+1/2} n^{-1/2}\big\}$ of the MISE, which replaces (15) in this case. The rates follow as stated.

Finally, $\beta = 5/2$, the integral in the previous case is bounded above by the integral of the function $f(t) = t^{-1}$ on $[-\omega_n, \omega_n]$. Hence, the upper bound of the

MISE is $\max\{\omega_n^{-4}\ln(\omega_n)\,,\,(\ln n)^{\beta+1/2}n^{-1/2}\}$. That proves the theorem in this case, too.

**Proof of Theorem 3.** Since $\sigma_n^2=\sigma_0^2$, the sets $\mathcal{G}_n$ in Lemma 1 are equal to the set $\mathcal{G}$ in (2). So, in view of Lemmas 1 and Lemma 2, the bounding techniques in the proof of Theorem 2 can be adopted respecting the changed sequences $(\sigma_n^2)_n$ and $(k_n)_n$.

**Proof of Theorem 4.** We construct functional sequences $(f_n)_n$, $(\tilde{f}_n)_n$, $(g_n)_n$, $(\tilde{g}_n)_n$ that satisfy the conditions of Lemma 3. Take

$$\varphi(t)=\begin{cases}\exp(a|t|), & |t|\le T,\\ C_2|t|^{-\beta}, & |t|>T,\end{cases}\tag{16}$$

with $a=(1/T)\ln(C_2T^{-\beta})<0$ with respect to the technical condition $(T/e)^\beta\ge C_2$. We easily check the conditions of Polya's criterion and recognize that $\varphi$ is the Fourier transform of a probability density $f$. Now $\varphi\in L_2(\mathbb{R})$ implies $f\in L_2(\mathbb{R})$ by Parseval's identity. Choose $\sigma^2\in(0,\sigma_0^2)$ and a sequence $(\sigma_n^2)_n$ in $(0,\sigma_0^2)$ with $\sigma_n^2\downarrow\sigma^2$. Take

$$\varphi_n(t)=\begin{cases}\varphi(t)\exp\left(\frac{1}{2}(\sigma^2-\sigma_n^2)t^2\right), & |t|\le t_n,\\ \varphi(t)\exp\left(\frac{1}{2}(\sigma^2-\sigma_n^2)t_n^2\right), & |t|>t_n,\end{cases}\tag{17}$$

with $t_n^2=\big(2/(\sigma_n^2-\sigma^2)\big)\ln(C_2/C_1)>0$. This implies $t_n\underset{n\to\infty}{\longrightarrow}+\infty$. The conditions of Polya's criterion with respect to $\varphi_n$, as well as the membership of $\varphi_n$ in $L_2(\mathbb{R})$, can also be verified so that $\varphi_n$ is the Fourier transform of a square integrable density $\overline{f}_n$. Since the defining condition of $\mathcal{F}$ is also satisfied by $\overline{f}_n$ as well as $f$, membership of $f$ and $\overline{f}_n$ in $\mathcal{F}$ follows. In Lemma 3, the functional sequences $(\tilde{f}_n)_n$, $(f_n)_n$ are specified by $f_n=f$ and $\tilde{f}_n=\overline{f}_n$. The sequences $(g_n)_n$ and $(\tilde{g}_n)_n$ in Lemma 3 are $g_n=g_{\sigma_n^2}=\mathrm{N}(\mu,\sigma_n^2)$ and $\tilde{g}_n=g_{\sigma^2}=\mathrm{N}(\mu,\sigma^2)$. This implies $\psi_{h_n}(t)=\psi_f(t)\psi_{g_{\sigma_n^2}}(t)=\varphi(t)\exp(i\mu t)\exp\big(-(1/2)\sigma_n^2t^2\big)$ and

$$\psi_{\tilde{h}_n}(t)=\psi_{\overline{f}_n}(t)\psi_{g_{\sigma^2}}(t)=\begin{cases}\varphi(t)\exp(i\mu t)\exp\left(-\frac{1}{2}\sigma_n^2t^2\right), & |t|\le t_n,\\ \frac{C_1}{C_2}\varphi(t)\exp(i\mu t)\exp\left(-\frac{1}{2}\sigma^2t^2\right), & |t|>t_n.\end{cases}$$

The $L_2(\mathbb{R})$-distance between $h_n$ and $\tilde{h}_n$ is bounded above by

$$\begin{aligned}\|h_n-\tilde{h}_n\|^2_{L_2(\mathbb{R})}&=(2\pi)^{-1}\|\psi_{h_n}-\psi_{\tilde{h}_n}\|^2_{L_2(\mathbb{R})}=(2\pi)^{-1}\int_{|t|>t_n}|\psi_{h_n}(t)-\psi_{\tilde{h}_n}(t)|^2dt\\ &\le\frac{4}{\pi}\|\varphi\|^2_{L_2(\mathbb{R})}\exp(-\sigma^2t_n^2)\\ &\le\frac{4}{\pi}\|\varphi\|^2_{L_2(\mathbb{R})}\exp\Big(-2\sigma^2\big(\ln(C_2)-\ln(C_1)\big)(\sigma_n^2-\sigma^2)^{-1}\Big),\end{aligned}\tag{18}$$

due to $\sigma_n^2 > \sigma^2$.

The function $\psi_{h_n}(t) = \varphi(t)\exp(i\mu t - 0.5\sigma_n^2 t^2)$ is continuous and differentiable on $\mathbb{R}\backslash\{-T,0,T\}$ with derivative $\varphi_n'(t) = \big(\varphi'(t) + (i\mu - \sigma_n^2 t)\varphi(t)\big)\exp\big(i\mu t - (1/2)\sigma_n^2 t^2\big)$. Due to the boundedness and the exponential-type decay of the derivative, one can see that $\varphi_n'$ is square integrable, so $\varphi_n \in W^1$. The $L_2(\mathbb{R})$-norm of $\varphi_n'$ is uniformly bounded above by

$$\begin{aligned}\|\psi_{h_n}'\|_{L_2(\mathbb{R})}^2 &= \int \big|\big(\varphi'(t) + (i\mu - \sigma_n^2 t)\varphi(t)\big)\exp\big(i\mu t - \frac{1}{2}\sigma_n^2 t^2\big)\big|^2 dt \\ &\le \int \exp(-\sigma^2 t^2)\big(|\varphi'(t)| + (|\mu| + \sigma_0^2|t|)|\varphi(t)|\big)^2 dt < +\infty.\end{aligned}$$

Now we concentrate on $\psi_{\tilde{h}_n}$. For this function we also have continuity and differentiability on $\mathbb{R}\backslash\{-t_n,-T,0,T,t_n\}$. The weak derivative is given by

$$\psi_{\tilde{h}_n}'(t) = \begin{cases} \big(\varphi'(t) + (i\mu - \sigma_n^2 t)\varphi(t)\big)\exp\big(i\mu t - \frac{1}{2}\sigma_n^2 t^2\big), & |t| < t_n, \\ \big(\varphi'(t) + (i\mu - \sigma^2 t)\varphi(t)\big)\exp\big(i\mu t - \frac{1}{2}\sigma^2 t^2\big)\frac{C_1}{C_2}, & |t| > t_n. \end{cases}$$

The $L_2(\mathbb{R})$-norm possesses an upper bound that is independent of $n$:

$$\begin{aligned}\|\psi_{\tilde{h}_n}'\|_{L_2(\mathbb{R})}^2 &\le \int \Big|\big(\varphi'(t) + (i\mu - \sigma_n^2 t)\varphi(t)\big)\exp\big(i\mu t - \frac{1}{2}\sigma_n^2 t^2\big)\Big|^2 dt \\ &\quad + \frac{C_1}{C_2}\int \Big|\big(\varphi'(t) + (i\mu - \sigma^2 t)\varphi(t)\big)\exp\big(i\mu t - \frac{1}{2}\sigma^2 t^2\big)\Big|^2 dt \\ &\le 2\int \big(|\varphi'(t)| + (|\mu| + \sigma_0^2|t|)\varphi(t)\big)^2 \exp(-\sigma^2 t^2)dt < \infty.\end{aligned}$$

Now we apply Lemma 3. We set $R_n = n^5$ and $\sigma_n^2 = \sigma^2 + d^{-1}(\ln n)^{-1}$ with $d = 19/\big(4\sigma^2\ln(C_2/C_1)\big) + 1$, which implies $t_n = \sqrt{2\ln(C_2/C_1)d}(\ln n)^{0.5}$; then (6) becomes $\big(n^{-\sigma^2\ln(C_2/C_1)} + \mathit{const}.n^{-1/4}\big)\|f_n - f\|_{L_2(\mathbb{R})}^{-2} \underset{n\to\infty}{\longrightarrow} 0$. It remains to search for a lower bound of $\|f_n - f\|_{L_2(\mathbb{R})}^2$. Let $(m_n)_n$ be a sequence satisfying $T \le m_n \le t_n$ for almost all $n$. Then we get

$$\begin{aligned}\|f_n - f\|_{L_2(\mathbb{R})}^2 &= (2\pi)^{-1}\|\varphi_n - \varphi\|_{L_2(\mathbb{R})}^2 \\ &\ge (2\pi)^{-1}\int_{m_n}^{t_n} |\varphi(t)|^2\big|\exp\big(\frac{1}{2}(\sigma^2 - \sigma_n^2)t^2\big) - 1\big|^2 dt \\ &\ge \mathit{const}.\big(m_n^{1-2\beta} - t_n^{1-2\beta}\big)\Big(\exp\big(\frac{1}{2}(\sigma^2 - \sigma_n^2)m_n^2\big) - 1\Big)^2.\end{aligned}$$

Consider the case $1 < \beta < 2.5$, for which we define the sequence $m_n = 0.5\sqrt{2\ln(C_2/C_1)d}(\ln n)^{0.5}$. Hence,

$$\|f_n - f\|_{L_2(\mathbb{R})}^2 \ge \mathit{const}.\Big(2\big(\ln(C_2) - \ln(C_1)\big)d\Big)^{0.5-\beta}\big(2^{2\beta-1} - 1\big)(\ln n)^{0.5-\beta}.$$

Then (6) is valid and the MISE is bounded below by $const.(\ln n)^{0.5-\beta}$.

Now consider the case $\beta \geq 2.5$ with $m_n = T$. Then $\|f_n - f\|^2_{L_2(\mathbb{R})} \geq const.(\exp((1/2)(\sigma^2 - \sigma_n^2)T^2) - 1)^2 \geq const.(\sigma^2 - \sigma_n^2)^2 \geq const.(\ln n)^{-2}$. The theorem follows.

## 4. Simulations

The estimation problem (2) is considered under several conditions, based on $n = 1,000$ observations. In Figures 1−3 the target density, which is plotted with dotted lines, is the standard Laplace density convolved with itself (parameters $C_2 = 1$, $C_1 = 1/4$, $T = 1$, $\beta = 4$), while in Figures 4−6, the three times self-convolved Laplace density is the density of interest (parameters $C_2 = 1$, $C_1 = 1/8$, $T = 1$, $\beta = 6$). In Figures 1 and 4, the error density is $N(0, 0.5)$ and the known upper bound is $\sigma_0^2 = 1$, while in Figures 2, 3, 4, 5 and 6, $N(0, 4)$ is the error density with $\sigma_0^2 = 8$. In Figures 3 and 6, the robustness referring to misspecification of $C_1$ and $\beta$ is illustrated; we have used $C_1' = 1/3$, $\beta' = 5$ in Figure 3 and $C_1' = 0.1$, $\beta' = 5$ in Figure 6. We show three independent replications of our estimator (5) in each figure, plotted with a solid line. The estimated variances are given below the figures. Furthermore, the classical deconvolution estimator, which uses the real error variance in its construction, is plotted with a dashed line.

We see that, in Figures 1 and 4, the trunction of the variance estimator is significant. Also, we observe only a slight loss of quality due to misspecification in Figures 3 and 6. Numerical errors and the non-negligible constants occurring in the slow rates of convergence make a discussion of our theoretical results based on the simulations difficult. However, simulations show that our estimator performs well in practical applications, and it shows only slight deterioration compared to the classical deconvolution estimator in the situation of a known variance.

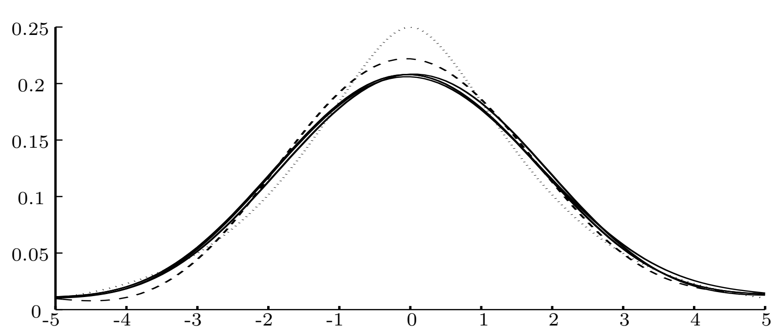


Figure 1. $\hat{\sigma_n^2} = 0; 0; 0.12.$

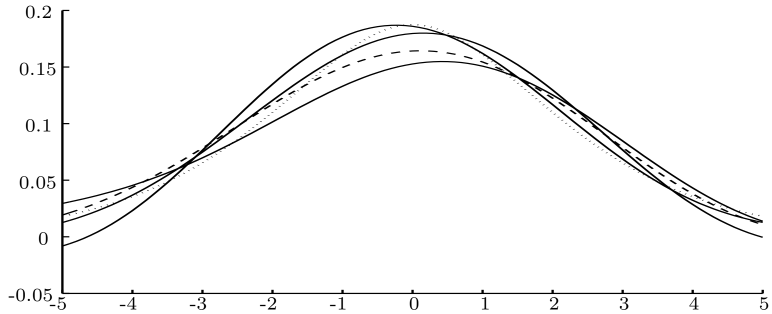


Figure 2. $\hat{\sigma_n^2} = 2.70; 5.14; 4.40.$

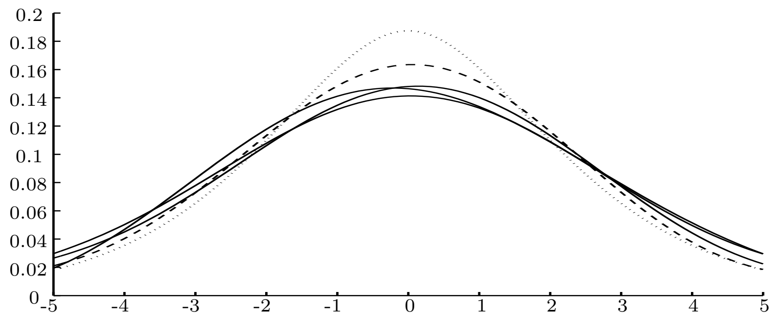


Figure 3. $\hat{\sigma_n^2} = 2.59; 4.32; 3.12.$

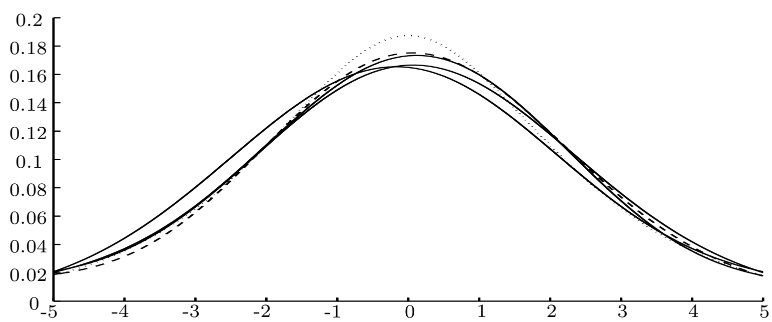


Figure 4. $\hat{\sigma_n^2} = 0; 0; 0.$

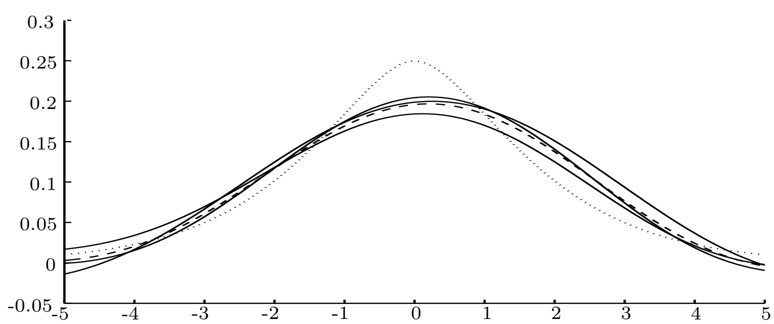


Figure 5. $\hat{\sigma_n^2} = 2.65; 5.19; 5.10.$

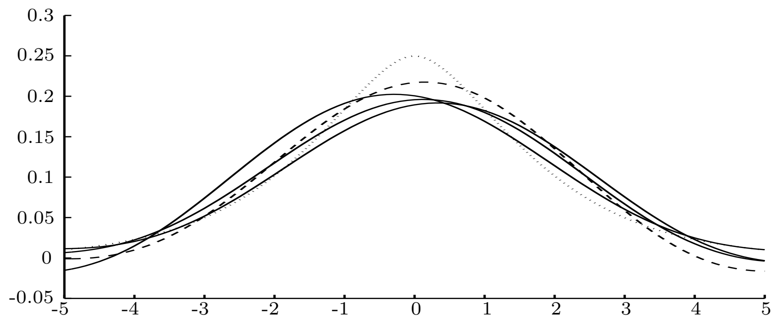


Figure 6. $\hat{\sigma_n^2} = 1.00; 2.52; 2.09$

**Acknowledgement**

The author is grateful to the referee for some constructive comments on this note.

## References

Butucea, C. and Matias, C. (2004). Minimax estimation of the noise level and of the deconvolution density in a semiparametric convolution model. To appear in *Bernoulli*.

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

Carroll, R. J. and Hall, P. (2004). Low order approximations in deconvolution and regression with errors in variables. *J. Roy. Statist. Soc. Ser. B* **66**, 31-46.

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

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

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

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

Durrett, R. (1996). *Probability: theory and examples*. 2nd edition. Duxberry Press.

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

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

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

Hesse, C. H. and Meister, A. (2004). Optimal iterative density deconvolution. *J.J. Nonparametr. Stat.* **16**, 879-900.

Hesse, C. H. (1999). Data-driven deconvolution. *Journal of Nonparametric Statistics.* **10**, 343-373.

Liu, M. C. and Taylor, R. C. (1989). A consistent nonparametric density estimator for the deconvolution problem. *Canad. J. Statist.* **17**, 427-438.

Matias, C. (2002). Semiparametric deconvolution with unknown noise variance. *ESAIM, Probab. Stat.* **6**, 271-292.

Meister, A. (2003). Robustheitseigenschaften von Dekonvolutionsdichteschätzern bezüglich Missspezifikation der Fehlerdichte. Dissertation, University of Stuttgart, Germany.

Meister, A. (2004a). Deconvolution density estimation with a testing procedure for the error distribution. Technical Report, University of Stuttgart, Germany.

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

Neumann, M. H. (1997). On the effect of estimating the error density in nonparametric deconvolution. *J. Nonparametr. Stat.* **7**, 307-330.

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

Institut für Stochastik und Anwendungen, Universität Stuttgart, Pfaffenwaldring 57, D-70569 Stuttgart, Germany.

E-mail: meistear@mathematik.uni-stuttgart.de

(Received May 2004; accepted May 2005)