Universität Stuttgart

Fachbereich Mathematik

Deconvolution Density Estimation with a Testing Procedure for the Error Distribution

Alexander Meister

Stuttgarter Mathematische Berichte

Preprint 2004/004

Universität Stuttgart
Fachbereich Mathematik

Deconvolution Density Estimation with a Testing Procedure for the Error Distribution

Alexander Meister

Stuttgarter Mathematische Berichte

Preprint 2004/004

Fachbereich Mathematik
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 basic task in deconvolution density estimation is the estimation of a probability density f based on contaminated observations $Y_1, \dots, Y_n$. Mathematically spoken, identically distributed random variables $X_1, \dots, X_n$ with density f and also identically distributed random variables $\varepsilon_1, \dots, \varepsilon_n$ with density g which represent the error or the contamination are given. The random variables $X_1, \varepsilon_1, \dots, X_n, \varepsilon_n$ are independent. The random variables $X_1, \dots, X_n$ whose density f shall be estimated cannot be observed directly but only the contaminated data $Y_1, \dots, Y_n$ defined by

$$Y_j = X_j + \varepsilon_j , \qquad \forall j \in \{1, \dots, n\}$$

can be used for constructing the estimator. So the density h of Y_j equals $h = f * g$, i.e. the convolution of the densities f and g. Now, it is aimed to find an estimator $\hat{f}_n$ of the density f based on $Y_1, \dots, Y_n$. Deterministic a-priori knowledge of nonparametric character about the density f is usually given, for example conditions referring to the asymptotic behaviour of the Fourier transform. These informations are expressed by defining a density class $\mathcal{F}$ and stipulating $f \in \mathcal{F}$. In the classical approach, the error density g is assumed to be exactly known (see for example Devroye (1989), Fan (1991), Fan (1993), Hesse (1999), Liu and Taylor (1990), Stefanski and Carroll (1990)). Therefore g can be used in the estimator's construction.
However, in lots of practical situations the assumption of a perfectly known error density is not realistic. So, in the papers of Efromovich (1997) and Neumann (1997), the error density is assumed to be unknown but can be estimated by additional empirical data which directly hail from the error distribution.
In this paper, I regard the situation that the error density is neither exactly known nor can be estimated by further observations. Therefore, I introduce the set $\mathcal{G}$ of all densities that can possibly be the error density and the a-priori knowledge about the error density can be expressed by the stipulation $g \in \mathcal{G}$. The case of a known error density is inbedded and equivalent with $\#\mathcal{G} = 1$. So it is aimed to achieve consistency although g is unknown. Therefore, I define the expression of robust consistency. An estimator $(\hat{f}_n)_n$ is called robustly d^k-consistent if

$$d(\hat{f}_n(Y_1, \dots, Y_n), f)^k \stackrel{p}{\longrightarrow} 0$$

is valid for all $f \in \mathcal{F}$ and for all $g \in \mathcal{G}$ as $n \to \infty$. The symbol $\stackrel{p}{\longrightarrow}$ means weak convergence. d is an arbitrary metric with a coordinated metrical space X and $k > 0$. An estimator $(\hat{f}_n)_n$ is called uniformly robustly d^k-consistent if

$$\sup_{f \in \mathcal{F}} \sup_{g \in \mathcal{G}} E_{f*g}\, d(\hat{f}_n(Y_1, \dots, Y_n), f)^k \stackrel{n \to \infty}{\longrightarrow} 0.$$

This is a stronger version of consistency. In order to construct a robustly consistent estimator, we have to do some theoretical work for finding conditions to the density classes $\mathcal{F}$ and $\mathcal{G}$, under which such an estimation is impossible. This is done in section 2. Of course, negative results are not very exciting but they help to avoid searching for non-existing robustly consistent estimators. In section 3, I regard specially defined density classes $\mathcal{F}$ and $\mathcal{G}$, for which a uniformly robustly $L_2(\mathbb{R})^2$-consistent estimator can indeed be constructed and calculated. After this work had been completed, I noticed that Butucea and Matias (2003) considered a similar deconvolution procedure with supersmooth error densities with unknown scaling parameter. I considered the situation of two possible error densities – a smooth versus a supersmooth density. While Butucea and Matias regarded pointwise consistency, I considered the MISE (=mean integrated square error) of the estimator. Then, I show that these rates are optimal under the given conditions. In section 4, I give the proofs of the lemmas and theorems.

2 Overlapping of $\mathcal{F}$ and $\mathcal{G}$

First, I define the meaning of overlapping density sets in this work: We call the density classes $\mathcal{F}$ and $\mathcal{G}$ overlapping if

$$\exists f, \tilde{f} \in \mathcal{F} \, \exists g, \tilde{g} \in \mathcal{G} \, : \, f \neq \tilde{f} \, \wedge \, f * g = \tilde{f} * \tilde{g}.$$

Hence, we can derive the first theorem

Theorem 1 *Assume the density classes $\mathcal{F}$ and $\mathcal{G}$ are overlapping. Let (X, d) be an arbitrary metrical space, let $\mathcal{F}$ be a subset of X. Then a robustly d^k-consistent estimator for any $k > 0$ does not exist.*

By this, we recognize the great meaning of overlapping in this case. It follows naturally that a uniformly robustly d^k-consistent estimator cannot exist, too. We look at some examples of overlapping density sets:

(1) The set $\mathcal{F} \cap \mathcal{G}$ consists of more than one element. Then we have two different $f, g \in \mathcal{F} \cap \mathcal{G}$ fulfilling

$$\underbrace{f}_{\in \mathcal{F}} * \underbrace{g}_{\in \mathcal{G}} = \underbrace{g}_{\in \mathcal{F}} * \underbrace{f}_{\in \mathcal{G}}.$$

So $\mathcal{F}$ and $\mathcal{G}$ are overlapping. This is quite intuitive. In the first case, f occurs as the estimated density and g as the error density and, in the second case, exactly the other way round. So in both situations the density of the observed data is $h = f * g$ and it cannot be decided whether f or g is the density which shall be estimated.

(2) We postulate the existence of $a \neq 0$, $f \in \mathcal{F}$ and $g \in \mathcal{G}$ so that

$$f(\cdot - a) \in \mathcal{F} \text{ and } g(\cdot + a) \in \mathcal{G}.$$

Then, for every $x \in \mathbb{R}$, we have

$$f(\cdot - a) * g(\cdot + a) \; = \; f * g$$

and $f \neq f(\cdot - a)$. Just imagine $f = f(\cdot - a)$. Then, by Fourier transformation, we receive $\psi_f(t) = \exp(ita) \cdot \psi_f(t)$ for all $t \in \mathbb{R}$. So, for every $t \in \mathbb{R}$, we have $\psi_f(t) = 0$ or $\exp(ita) = 1$. As $a \neq 0$, there is an interval of positive length with center 0 so that $\psi_f(t) = 0$ for all t in this interval except 0. Since $\psi_f(0) = 1$ holds, ψ_f is not continuous and we receive a contradiction as f is a density.

(3) Due to the importance of the following set of overlapping densities, I formulate

Lemma 1 *Let $\mathcal{H} := \{f * g \mid f \in \mathcal{F}, g \in \mathcal{G}\}$. Under the conditions $\mathcal{H} \subseteq \mathcal{F}$ and $\exists f \in \mathcal{F} : \psi_f(t) \neq 0 \, , \forall t \in \mathbb{R}$, the stipulation $\#\mathcal{G} \geq 2$ suffices to ensure the overlapping of $\mathcal{F}$ and $\mathcal{G}$.*

So, if the conditions referring to $\mathcal{F}$ in lemma 1 are fulfilled, the membership of more than one density in $\mathcal{G}$, which is equivalent with any kind of imperfect knowledge of the error density, makes a consistent estimation of f impossible. Lots of widely used density classes for $\mathcal{F}$ are affected by these conditions.

Theorem 2 *Let $\mathcal{F}$ be one of the sets*

$$\mathcal{F}^{(1)}_{(C,\beta)} := \{f \text{ density} \mid |\psi_f(t)| \leq C|t|^{-\beta} \, , \forall t \in \mathbb{R}\} \quad (\beta > 1, C > 0)$$
$$\mathcal{F}^{(2)}_{(C,\beta)} := \{f \text{ density} \mid \textstyle\int |\psi_f(t)|^2 (1 + |t|)^{2\beta} dt \, \leq C\} \quad (\beta \in \mathbb{R}, C > 0)$$

and assume that $\mathcal{G}$ consists of more than one element. Then $\mathcal{F}$ and $\mathcal{G}$ are overlapping.

$\mathcal{F}^{(1)}_{(C,\beta)}$ is used in Hesse and Meister (2001). $\mathcal{F}^{(2)}_{(C,\beta)}$ appears in Neumann (1997).

3 Uniformly Robustly $L_2(\mathbb{R})^2$-consistent estimation

Now, we consider the situation

$$\begin{aligned} \mathcal{F} &:= \{f \text{ density} \mid C_2 t^{-2} \geq |\psi_f(t)| \geq C_1 t^{-2}\,, \forall t \text{ with } |t| \geq T \geq 1\} \\ \mathcal{G} &:= \{g_L, g_N\}, \end{aligned} \tag{3.1}$$

with g_L being the Laplace density (i.e. $g_L(t) = \frac{1}{2}\exp(-|x|)$) and g_N being the standard normal distribution density. So $\mathcal{G}$ contains a smooth and a supersmooth density. T is fixed and known. The empirical data can be used twice: in order to estimate both the error density and, afterwards, the density f. One can recognize that for all $h = f * g_L$, $f \in \mathcal{F}$

$$|\psi_h(t)| = |\psi_{g_L}(t)| \cdot |\psi_f(t)| \geq \frac{C_1}{t^2} \cdot \frac{1}{1+t^2} = \frac{C_1}{t^2 + t^4} =: O(t)$$

for all $|t| \geq T$ and that for all $h = f * g_N$, $f \in \mathcal{F}$

$$|\psi_h(t)| = |\psi_{g_N}(t)| \cdot |\psi_f(t)| \leq \frac{C_2}{t^2} \cdot \exp(-t^2/2) =: U(t)$$

for all $|t| \geq T$. One can easily see by the definition of the functions O and U that $O(t_0) > U(t_0)$ is valid if $t_0 > T$ is large enough (for example $t_0 = 3$ if $T = 1, C_1 = 1/2, C_2 = 1$). By using the empirical Fourier transform, one can construct an estimator of the error density

$$\hat{g}_{t_0,n}(Y_1, \ldots, Y_n) := \begin{cases} g_L & \text{if } \left| \frac{1}{n} \sum\limits_{j=1}^{n} e^{it_0 Y_j} \right| \geq \frac{1}{2}(O(t_0) + U(t_0)) \\ \\ g_N & \text{if } \left| \frac{1}{n} \sum\limits_{j=1}^{n} e^{it_0 Y_j} \right| < \frac{1}{2}(O(t_0) + U(t_0)) \end{cases} \tag{3.2}$$

This estimator's risk can be bounded

Lemma 2

$$\begin{aligned} \sup_{f \in \mathcal{F}} P_{g_N * f}\left(\hat{g}_{t_0,n} = g_L\right) &\leq \tfrac{4(O(t_0) - U(t_0))^{-2}}{n}, \\ \sup_{f \in \mathcal{F}} P_{g_L * f}\left(\hat{g}_{t_0,n} = g_N\right) &\leq \tfrac{4(O(t_0) - U(t_0))^{-2}}{n}, \end{aligned}$$

So one can derive an estimator of the density f

$$\hat{f}_n(x) = \frac{1}{2\pi} \int\limits_{-\omega_n}^{\omega_n} \exp(-itx) \frac{1}{n} \sum_{j=1}^{n} \exp(itY_j)/\psi_{\hat{g}_{t_0,n}}(t)\, dt. \tag{3.3}$$

Hence, one receives the following result

Theorem 3 *The estimator defined by (3.3) is uniformly robustly $L_2(\mathbb{R})^2$-consistent in the problem (3.1). The rate of consistency in the case of the bandwidth selection $(\omega_n)_n = (\frac{1}{2}\sqrt{\ln n})_n$ is upper bounded by*

$$\sup_{g \in \mathcal{G}} \sup_{f \in \mathcal{F}} E_{f*g} \|\hat{f}_n - f\|^2_{L_2(\mathbb{R})} \leq Const. \cdot (\ln n)^{-3/2}$$

This rate of convergence of the MISE is optimal as the following theorem says.

Theorem 4 *Let $\hat{f}_n$ be an arbitrary estimator of f based on the contaminated data $Y_1, \ldots, Y_n$ in the estimation problem (3.1). The technical condition $T^2 \geq 3C_2$ is stipulated to hold. Then there is a $c > 0$ so that*

$$\sup_{g \in \mathcal{G}} \sup_{f \in \mathcal{F}} E_{f*g} \|\hat{f}_n - f\|^2_{L_2(\mathbb{R})} \geq c \cdot (\ln n)^{-3/2}$$

Notice that this lower bound corresponds to the lower bound in the case of an exactly known standard normal density g. Since only two densities can occur as the error distribution, a deterioration of the asymptotical quality of the estimation does not occur. Lower bounds with known error density have been studied in several papers but the class $\mathcal{F}$ in (3.1) has not yet been regarded as far as I know.

4 Proofs

Proof of theorem 1: We assume $\hat{f}_n$ to be a robustly d^k-consistent estimator with d being an arbitrary metric and $k > 0$, i.e. for any $\epsilon > 0$,

$$P_{f*g}\big(d(\hat{f}_n(Y_1,\ldots,Y_n),f)^k \,\geq\, \epsilon\big) \overset{n\to\infty}{\longrightarrow} 0\,, \forall f \in \mathcal{F}, g \in \mathcal{G}$$

holds. As $\mathcal{F}$ and $\mathcal{G}$ are overlapping, there are $f, \tilde{f} \in \mathcal{F}$ and $g, \tilde{g} \in \mathcal{G}$ with $f \neq \tilde{f}$ and $f*g = \tilde{f}*\tilde{g}$. Setting $\epsilon := \frac{1}{4}\min\{2^{1-k},1\}\cdot d(f,\tilde{f})^k$, the postulation $\epsilon > 0$ is fulfilled because of the positive definiteness of d. Then,

$$\begin{aligned}
&P_{f*g}\big(d(\hat{f}_n(Y_1,\ldots,Y_n),f)^k \geq \epsilon\big) \,+\, P_{\tilde{f}*\tilde{g}}\big(d(\hat{f}_n(Y_1,\ldots,Y_n),\tilde{f})^k \geq \epsilon\big) \overset{n\to\infty}{\longrightarrow} 0\\
&\geq\, P_{h=f*g=\tilde{f}*\tilde{g}}\big(d(\hat{f}_n(Y_1,\ldots,Y_n),f)^k \geq \epsilon \,\vee\, d(\hat{f}_n(Y_1,\ldots,Y_n),\tilde{f})^k \geq \epsilon\big)\\
&\geq\, P_h\big(d(\hat{f}_n(Y_1,\ldots,Y_n),f)^k \,+\, d(\hat{f}_n(Y_1,\ldots,Y_n),\tilde{f})^k \,\geq\, 2\epsilon\big)\\
&\geq\, P_h\Big(\min\{2^{1-k},1\}\big(d(\hat{f}_n(Y_1,\ldots,Y_n),f) \,+\, d(\hat{f}_n(Y_1,\ldots,Y_n),\tilde{f})\big)^k \,\geq\, 2\epsilon\Big)\\
&\geq\, P_h(4\epsilon \geq 2\epsilon)\\
&=\, 1.
\end{aligned}$$

So we receive a contradiction referring to the assumption of the existence of a robustly d^k-consistent estimator. ■

Proof of lemma 1: Choose $f \in \mathcal{F}$ so that $\psi_f(t) \neq 0$ holds for all $t \in \mathbb{R}$ and select $g, \tilde{g} \in \mathcal{G}$ with $g \neq \tilde{g}$. Hence, we have

$$\begin{aligned}
g*f &\in \mathcal{H} \subseteq \mathcal{F}\\
\tilde{g}*f &\in \mathcal{H} \subseteq \mathcal{F}
\end{aligned}$$

Using the commutation law of convolution, we get

$$\Rightarrow\ \underbrace{(g*f)}_{\in\mathcal{F}} * \underbrace{\tilde{g}}_{\in\mathcal{G}} \;=\; \underbrace{g}_{\in\mathcal{G}} * \underbrace{(f*\tilde{g})}_{\in\mathcal{F}}.$$

So it remains to be shown that $g*f \neq f*\tilde{g}$. Assume that $g*f = f*\tilde{g}$ holds. Hence, by Fourier transformation, we have $\psi_g\cdot\psi_f = \psi_{\tilde{g}}\cdot\psi_f$. As ψ_f vanishes nowhere by stipulation, it follows that $\psi_g = \psi_{\tilde{g}}$ and hence $g = \tilde{g}$ and we have a contradiction related to the assumption above. ■

Proof of theorem 2: The stipulations of lemma 1 have to be proven. Firstly, we regard $\mathcal{F} = \mathcal{F}^{(1)}_{(C,\beta)}$. Therefore, we consider the normal distribution density with mean 0 and variance σ^2 called f_{σ^2}. It is well-known that the Fourier transform

$$\psi_{f_{\sigma^2}}(t) \,=\, \exp\big(-\frac{1}{2}\sigma^2t^2\big)$$

vanishes nowhere for all $\sigma > 0$. So one has to choose $\sigma > 0$ appropriately so that f_{σ^2} is a member of $\mathcal{F}^{(1)}_{(C,\beta)}$. This is equivalent with

$$\exp(-\frac{1}{2}\sigma^2 t^2)|t|^\beta \leq C\,,\forall t.$$

By elementary analytic calculation, one receives that the function on the left side of the equation above possesses its maximum for $t = \sqrt{\beta}/\sigma$. Hence, the function is upper bounded by $\sqrt{\beta}^\beta \sigma^{-\beta} \exp(-\beta/2)$ and the equation is equivalent with

$$\sigma \geq C^{-1/\beta} e^{-1/2} \beta^{1/2}.$$

Under this condition, f_{σ^2} is in $\mathcal{F}^{(1)}_{(C,\beta)}$. Secondly, we regard $\mathcal{F} = \mathcal{F}^{(2)}_{(C,\beta)}$.

$$\int \exp(-\sigma^2 t^2)(1+|t|)^{2\beta}\,dt \overset{!}{\leq} C$$

with $\sigma > \sigma_0 > 0$. So $\exp(-\sigma_0^2 \cdot^2)(1+|\cdot|)^{2\beta}$ is an integrable upper bound. Furthermore, for all $t \neq 0$ and hence for (Lebesgue-)almost all $t \in \mathbb{R}$, we have

$$\lim_{\sigma\to\infty}\left(\exp(-\sigma^2 t^2)(1+|t|)^{2\beta}\right) = 0.$$

Using dominated convergence, it follows that

$$\int \exp(-\sigma^2 t^2)(1+|t|)^{2\beta}\,dt \overset{\sigma\to\infty}{\longrightarrow} 0.$$

So, for each $\beta > 1$, σ can be chosen sufficiently large so that

$$\int \exp(-\sigma^2 t^2)(1+|t|)^{2\beta}\,dt \leq C$$

and hence $f_{\sigma^2} \in \mathcal{F}^{(2)}_{(C,\beta)}$.

It remains to be shown that $\mathcal{H} \subseteq \mathcal{F}$ holds for both cases $\mathcal{F} = \mathcal{F}^{(1)}_{(C,\beta)}$ and $\mathcal{F} = \mathcal{F}^{(2)}_{(C,\beta)}$. Since every $g \in \mathcal{G}$ is a density and hence $|\psi_g(t)| \leq 1$ is valid for all $t \in \mathbb{R}$, it follows that for any $h \in \mathcal{H}$

$$|\psi_h(t)| = |\psi_f(t)| \cdot |\psi_g(t)| \leq |\psi_f(t)|$$

holds for all $t \in \mathbb{R}$. So we have

$$\begin{array}{l} |\psi_h(t)| \leq |\psi_f(t)| \leq C|t|^{-\beta}\,, \forall t \in \mathbb{R} \\ \int |\psi_h(t)|^2 (1+|t|)^{2\beta}\, dt \leq \int |\psi_f(t)|^2 (1+|t|)^{2\beta}\, dt \leq C. \end{array}$$

and the inclusion $\mathcal{H} \subseteq \mathcal{F}$ is valid in both cases. ■

Proof of lemma 2: proof of the first inequality:

$$
\begin{aligned}
&\sup_{f\in\mathcal{F}} P_{g_N*f}\left(\hat{g}_{t_0,n}=g_L\right)\\
&= \sup_{f\in\mathcal{F}} P_{g_N*f}\left(\left|\frac{1}{n}\sum_{j=1}^{n} e^{it_0Y_j}\right| \ \geq\ \frac{1}{2}(O(t_0)+U(t_0))\right)\\
&\leq \sup_{f\in\mathcal{F}} P_{g_N*f}\left(\left|\frac{1}{n}\sum_{j=1}^{n} e^{it_0Y_j}-\psi_{g_N*f}(t_0)\right| + \underbrace{|\psi_{g_N*f}(t_0)|}_{\leq U(t_0)} \ \geq\ \frac{1}{2}(O(t_0)+U(t_0))\right)\\
&\leq \sup_{f\in\mathcal{F}} P_{g_N*f}\left(\left|\frac{1}{n}\sum_{j=1}^{n} e^{it_0Y_j}-\psi_{g_N*f}(t_0)\right| \ \geq\ \frac{1}{2}(O(t_0)-U(t_0))\ >\ 0\right)\\
&\textit{(Markov inequality)}\\
&\leq \sup_{f\in\mathcal{F}} 4(O(t_0)-U(t_0))^{-2}\, var_{g_N*f}\left(\frac{1}{n}\sum_{j=1}^{n} e^{it_0Y_j}\right)\\
&= 4(O(t_0)-U(t_0))^{-2}\,\frac{1}{n}\sup_{f\in\mathcal{F}}\left(1-|\psi_{f*g_N}(t_0)|^2\right)\\
&\leq \frac{4(O(t_0)-U(t_0))^{-2}}{n}
\end{aligned}
$$

and the second one:

$$
\begin{aligned}
&\sup_{f\in\mathcal{F}} P_{h=g_L*f}\left(\hat{g}_{t_0,n}=g_N\right)\\
&= \sup_{f\in\mathcal{F}} P_{h=g_L*f}\left(\left|\frac{1}{n}\sum_{j=1}^{n} e^{it_0Y_j}\right| < \frac{O(t_0)+U(t_0)}{2}\right)\\
&= \sup_{f\in\mathcal{F}} P_{h=g_L*f}\left(\left|\frac{1}{n}\sum_{j=1}^{n} e^{it_0Y_j}\right| - |\psi_h(t_0)| < \frac{O(t_0)+U(t_0)}{2} - \underbrace{|\psi_h(t_0)|}_{\geq O(t_0)}\right)\\
&\leq \sup_{f\in\mathcal{F}} P_{h=g_L*f}\left(\left|\frac{1}{n}\sum_{j=1}^{n} e^{it_0Y_j}\right| - |\psi_h(t_0)| < \frac{U(t_0)-O(t_0)}{2} < 0\right)\\
&\leq \sup_{f\in\mathcal{F}} P_{h=g_L*f}\left(\left|\left|\frac{1}{n}\sum_{j=1}^{n} e^{it_0Y_j}\right| - |\psi_h(t_0)|\right| \ \geq\ \frac{O(t_0)-U(t_0)}{2}\right)\\
&\leq \sup_{f\in\mathcal{F}} P_{h=g_L*f}\left(\left|\frac{1}{n}\sum_{j=1}^{n} e^{it_0Y_j}-\psi_h(t_0)\right| \ \geq\ \frac{O(t_0)-U(t_0)}{2}\right)\\
&\leq \sup_{f\in\mathcal{F}} 4(O(t_0)-U(t_0))^{-2}\, var_{h=g_L*f}\left(\frac{1}{n}\sum_{j=1}^{n} e^{it_0Y_j}\right)\\
&\leq \frac{4(O(t_0)-U(t_0))^{-2}}{n}.
\end{aligned}
$$

■

Proof of theorem 3:

Using the Parseval identity, we have

$$
\begin{aligned}
&\sup_{g\in\mathcal{G}}\sup_{f\in\mathcal{F}} E_{f*g}\|\hat f_n - f\|^2_{L_2(\mathbb{R})}\\
&\le \sup_{f\in\mathcal{F}} E_{f*g_N}\|\hat f_n - f\|^2_{L_2(\mathbb{R})} + \sup_{f\in\mathcal{F}} E_{f*g_L}\|\hat f_n - f\|^2_{L_2(\mathbb{R})}\\
&= \tfrac{1}{2\pi}\sup_{f\in\mathcal{F}} E_{f*g_N}\int\left|\tfrac{\chi_{[-\omega_n,\omega_n]}}{\psi_{\hat g_{t_0,n}}(t)}\tfrac{1}{n}\sum_{j=1}^n e^{itY_j} - \psi_f(t)\right|^2 dt\\
&+ \tfrac{1}{2\pi}\sup_{f\in\mathcal{F}} E_{f*g_L}\int\left|\tfrac{\chi_{[-\omega_n,\omega_n]}}{\psi_{\hat g_{t_0,n}}(t)}\tfrac{1}{n}\sum_{j=1}^n e^{itY_j} - \psi_f(t)\right|^2 dt
\end{aligned}
$$

$$
\begin{aligned}
&\le \tfrac{1}{2\pi}\sup_{f\in\mathcal{F}} E_{f*g_N}\int 2\left|\tfrac{\chi_{[-\omega_n,\omega_n]}}{\psi_{\hat g_{t_0,n}}(t)}\tfrac{1}{n}\sum_{j=1}^n e^{itY_j} - \tfrac{\chi_{[-\omega_n,\omega_n]}}{\psi_{g_N}(t)}\tfrac{1}{n}\sum_{j=1}^n e^{itY_j}\right|^2 dt\\
&+ \tfrac{1}{2\pi}\sup_{f\in\mathcal{F}} E_{f*g_N}\int 2\left|\tfrac{\chi_{[-\omega_n,\omega_n]}}{\psi_{g_N}(t)}\tfrac{1}{n}\sum_{j=1}^n e^{itY_j} - \psi_f(t)\right|^2 dt\\
&+ \tfrac{1}{2\pi}\sup_{f\in\mathcal{F}} E_{f*g_L}\int 2\left|\tfrac{\chi_{[-\omega_n,\omega_n]}}{\psi_{\hat g_{t_0,n}}(t)}\tfrac{1}{n}\sum_{j=1}^n e^{itY_j} - \tfrac{\chi_{[-\omega_n,\omega_n]}}{\psi_{g_L}(t)}\tfrac{1}{n}\sum_{j=1}^n e^{itY_j}\right|^2 dt\\
&+ \tfrac{1}{2\pi}\sup_{f\in\mathcal{F}} E_{f*g_L}\int 2\left|\tfrac{\chi_{[-\omega_n,\omega_n]}}{\psi_{g_L}(t)}\tfrac{1}{n}\sum_{j=1}^n e^{itY_j} - \psi_f(t)\right|^2 dt\\
&\le \tfrac{1}{\pi}\sup_{f\in\mathcal{F}}\int_{-\omega_n}^{\omega_n} E_{f*g_N}\left(\underbrace{\left|\frac{1}{n}\sum_{j=1}^n e^{itY_j}\right|^2}_{\le 1}\left|\tfrac{1}{\psi_{\hat g_{t_0,N}(t)}} - \tfrac{1}{\psi_{g_N}(t)}\right|^2\right) dt\\
&+ \tfrac{1}{\pi}\sup_{f\in\mathcal{F}} E_{f*g_N}\int\left|\tfrac{\chi_{[-\omega_n,\omega_n]}}{\psi_{g_N}(t)}\tfrac{1}{n}\sum_{j=1}^n e^{itY_j} - \psi_f(t)\right|^2 dt\\
&+ \tfrac{1}{\pi}\sup_{f\in\mathcal{F}}\int_{-\omega_n}^{\omega_n} E_{f*g_L}\left(\underbrace{\left|\frac{1}{n}\sum_{j=1}^n e^{itY_j}\right|^2}_{\le 1}\left|\tfrac{1}{\psi_{\hat g_{t_0,N}(t)}} - \tfrac{1}{\psi_{g_L}(t)}\right|^2\right) dt\\
&+ \tfrac{1}{\pi}\sup_{f\in\mathcal{F}} E_{f*g_L}\int\left|\tfrac{\chi_{[-\omega_n,\omega_n]}}{\psi_{g_L}(t)}\tfrac{1}{n}\sum_{j=1}^n e^{itY_j} - \psi_f(t)\right|^2 dt\\
&\le \tfrac{1}{\pi}\sup_{f\in\mathcal{F}}\int_{-\omega_n}^{\omega_n}\left|\tfrac{1}{\psi_{g_L}(t)} - \tfrac{1}{\psi_{g_N}(t)}\right|^2 dt\cdot P_{h=g_N*f}\left(\hat g_{t_0,n} = g_L\right)\\
&+ \tfrac{1}{\pi}\sup_{f\in\mathcal{F}} E_{f*g_N}\int\left|\tfrac{\chi_{[-\omega_n,\omega_n]}}{\psi_{g_N}(t)}\tfrac{1}{n}\sum_{j=1}^n e^{itY_j} - \psi_f(t)\right|^2 dt\\
&+ \tfrac{1}{\pi}\sup_{f\in\mathcal{F}}\int_{-\omega_n}^{\omega_n}\left|\tfrac{1}{\psi_{g_L}(t)} - \tfrac{1}{\psi_{g_N}(t)}\right|^2 dt\cdot P_{h=g_L*f}\left(\hat g_{t_0,n} = g_N\right)\\
&+ \tfrac{1}{\pi}\sup_{f\in\mathcal{F}} E_{f*g_L}\int\left|\tfrac{\chi_{[-\omega_n,\omega_n]}}{\psi_{g_L}(t)}\tfrac{1}{n}\sum_{j=1}^n e^{itY_j} - \psi_f(t)\right|^2 dt
\end{aligned}
$$

(lemma 2)

$$\leq \frac{8(O(t_0)-U(t_0))^{-2}}{\pi n} \cdot \int_{-\omega_n}^{\omega_n} \left|\frac{1}{\psi_{g_L(t)}} - \frac{1}{\psi_{g_N}(t)}\right|^2 dt$$

$$+ \frac{1}{\pi} \sup_{f\in\mathcal{F}} E_{f*g_N} \int \left|\frac{\chi_{[-\omega_n,\omega_n]}}{\psi_{g_N}(t)} \frac{1}{n}\sum_{j=1}^{n} e^{itY_j} - \psi_f(t)\right|^2 dt$$

$$+ \frac{1}{\pi} \sup_{f\in\mathcal{F}} E_{f*g_L} \int \left|\frac{\chi_{[-\omega_n,\omega_n]}}{\psi_{g_L}(t)} \frac{1}{n}\sum_{j=1}^{n} e^{itY_j} - \psi_f(t)\right|^2 dt$$

$$\leq \frac{8(O(t_0)-U(t_0))^{-2}}{\pi n} \cdot \int_{-\omega_n}^{\omega_n} \left|\frac{1}{\psi_{g_L(t)}} - \frac{1}{\psi_{g_N}(t)}\right|^2 dt + \frac{4}{\pi} \sup_{f\in\mathcal{F}} \int_{\omega_n}^{\infty} |\psi_f(t)|^2 dt$$

$$+ \frac{1}{\pi} \sup_{f\in\mathcal{F}} \int_{-\omega_n}^{\omega_n} var_{f*g_N} \left(\frac{1}{n}\sum_{j=1}^{n} e^{itY_j}/\psi_{g_N}(t)\right) dt$$

$$+ \frac{1}{\pi} \sup_{f\in\mathcal{F}} \int_{-\omega_n}^{\omega_n} var_{f*g_L} \left(\frac{1}{n}\sum_{j=1}^{n} e^{itY_j}/\psi_{g_L}(t)\right) dt$$

$$\leq \frac{8(O(t_0)-U(t_0))^{-2}}{\pi n} \cdot \int_{-\omega_n}^{\omega_n} \left|\frac{1}{\psi_{g_L(t)}} - \frac{1}{\psi_{g_N}(t)}\right|^2 dt + \frac{4}{\pi} \sup_{f\in\mathcal{F}} \int_{\omega_n}^{\infty} |\psi_f(t)|^2 dt$$

$$+ \frac{2}{n\pi} \int_0^{\omega_n} |\psi_{g_N}(t)|^{-2} dt + \frac{2}{n\pi} \int_0^{\omega_n} |\psi_{g_L}(t)|^{-2} dt$$

$$\leq \frac{32(O(t_0)-U(t_0))^{-2}}{\pi n} \cdot \int_0^{\omega_n} \left(\left|\frac{1}{\psi_{g_L(t)}}\right|^2 + \left|\frac{1}{\psi_{g_N}(t)}\right|^2 dt\right) + \frac{4}{\pi} \sup_{f\in\mathcal{F}} \int_{\omega_n}^{\infty} |\psi_f(t)|^2 dt$$

$$+ \frac{2}{n\pi} \int_0^{\omega_n} |\psi_{g_N}(t)|^{-2} dt + \frac{2}{n\pi} \int_0^{\omega_n} |\psi_{g_L}(t)|^{-2} dt$$

$$\leq \frac{4}{\pi} \sup_{f\in\mathcal{F}} \int_{\omega_n}^{\infty} |\psi_f(t)|^2 dt + \left(\frac{2}{n\pi} + \frac{32(O(t_0)-U(t_0))^{-2}}{\pi n}\right) \int_0^{\omega_n} |\psi_{g_N}(t)|^{-2} dt$$

$$+ \left(\frac{2}{n\pi} + \frac{32(O(t_0)-U(t_0))^{-2}}{\pi n}\right) \int_0^{\omega_n} |\psi_{g_L}(t)|^{-2} dt$$

This term converges to 0 if and only if all three nonnegative summands do so. The rate of convergence equals the rate of the most slowly converging summand. The first summand can be upper bounded by

$$\int_{\omega_n}^{\infty} t^{-4} dt = \frac{1}{3}\omega_n^{-3},$$

the second one by

$$n^{-1}\omega_n \exp(\omega_n^2)$$

and the third one by

$$n^{-1}\omega_n(1+\omega_n^2)^2.$$

These bounds are given up to multiplication of a positive constant. If we optimize the rate of convergence by choosing the bandwidth sequence as given in the theorem, then we receive $(\ln n)^{-3/2}$ as rate of convergence of the complete term. ■

Proof of theorem 4:
First, I define the function

$$\varphi_n(t) := \begin{cases} \frac{T^{-2}C_2-1}{T}t+1 & \text{if } |t| \leq T \\ C_2 t^{-2} & \text{else.} \end{cases}$$

Let $(M_n)_n$ be a sequence with $M_n > T$ which will exactly be determined later. One can construct another function $\tilde{\varphi}_n$ which equals φ_n on the restriction to $[-M_n, M_n]$. I apply the tangent of φ_n in $t = M_n$ and $t = -M_n$ for $\tilde{\varphi}_n$ in the sections $[M_n, D_n]$ and $[-M_n, -D_n]$ with D_n denoting the unique intersection of the tangent and the curve $C_1\bullet^{-2}$ in $|t| \geq M_n$. For the sections $[D_n, +\infty)$ and $(-\infty, -D_n]$, $\tilde{\varphi_n}(t)$ equals $C_1 t^{-2}$. So, we have

$$\tilde{\varphi}_n \;=\; \begin{cases} \varphi_n(t) & \text{if } |t| \leq M_n \\ C_2 M_n^{-2} - 2C_2 M_n^{-3}(t - M_n) & \text{if } M_n < |t| \leq D_n \\ C_1 t^{-2} & \text{if } |t| > D_n. \end{cases}$$

Notice that, due to the definition of D_n, the equation

$$C_2 M_n^{-2} - 2C_2 M_n^{-3}(D_n - M_n) \;=\; C_1 D_n^{-2}$$

is valid. This is equivalent with

$$3C_2\left(\frac{D_n}{M_n}\right)^2 \;-\; 2C_2\left(\frac{D_n}{M_n}\right)^3 \;=\; C_1.$$

The polynomial function $P(x) := 3C_2x^2 - 2C_2x^3 - C_1$ decreases strictly monotoniously for $x > 1$ and tends to $-\infty$ if $x \to +\infty$. Combining this with $P(1) > 0$, it is evident that there is exactly one $D > 1$ fulfilling $P(D) = 0$. Hence, D_n equals

$$D_n \;=\; D \cdot M_n,$$

respecting that D does not depend on n. One can recognize that φ_n is the Fourier transform of a probability density by using Polya's criterion (see Durrett (1996)). The conditions are non-negativity, convexity, decreasingness, continuity of φ_n on $(0, \infty)$ and $\varphi_n(t) = \varphi_n(-t)$, $\varphi_n(0) = 1$ as well as $\lim_{t\uparrow\infty}\varphi_n(t) = 0$. Respecting the technical stipulation $T^2 \geq 3C_2$, these conditions can be proven for φ_n. Regarding the construction of $\tilde{\varphi}_n$, those properties can also be seen. So, φ_n is the Fourier transform of a probability density, too. Denoting these densities f_n and $\tilde{f}_n$, we have

$$\varphi_n \;=\; \psi_{f_n} \text{ and } \tilde{\varphi}_n \;=\; \psi_{\tilde{f}_n}.$$

One can also see that φ_n and $\tilde{\varphi}_n$ fulfill the upper and lower bound stipulations in the defining condition of $\mathcal{F}$. So, we receive

$$f_n, \tilde{f}_n \;\in\; \mathcal{F}.$$

Denoting $h_n = f_n * g_N$ and $\tilde{h}_n = \tilde{f}_n * g_N$, one recognizes that ψ_{h_n} is continuous on $\mathbb{R}$ and differentiable on $\mathbb{R}\backslash\{-T, T\}$. Its derivative is given by

$$\psi'_{h_n}(t) \;=\; \exp(-\frac{1}{2}t^2) \cdot \big(\varphi'_n(t) \;+\; \varphi_n(t) \cdot (-t)\big).$$

Since φ_n and

$$\varphi'_n(t) \;=\; \begin{cases} \frac{T^{-2}C_2 - 1}{T} & \text{if } |t| \leq T \\ -2C_2 t^{-3} & \text{else} \end{cases}$$

are uniformly upper bounded, $\|\psi'_{h_n}\|_{L_2(\mathbb{R})}$ possesses an upper bound which is independent of n. Considering $\psi_{\tilde{h}_n}$, one can derive continuity on $\mathbb{R}$ and differentiability on $\mathbb{R}\backslash\{-DM_n, -T, T, DM_n\}$. Accordingly, we have

$$\psi'_{\tilde{h}_n}(t) \;=\; \exp(-\frac{1}{2}t^2) \cdot \big(\tilde{\varphi}'_n(t) \;+\; \tilde{\varphi}_n(t) \cdot (-t)\big).$$

In analogy, one can see that $\tilde{\varphi}_n$ and

$$\tilde{\varphi}_n'(t) = \begin{cases} \varphi_n'(t) & \text{if } |t| \le M_n \\ -2C_2 M_n^{-3} & \text{if } M_n < |t| \le DM_n \\ -2C_1 t^{-3} & \text{if } |t| > DM_n, \end{cases}$$

are also uniformly upper bounded, respecting $M_n > T$. Hence, ψ_{h_n} and $\psi_{\tilde{h}_n}$ are weakly differentiable with the derivative being uniformly (relating to n) bounded in the $L_2(\mathbb{R})$-norm. So ψ_{h_n} and $\psi_{\tilde{h}_n}$ are members of the Sobolev space of order 1 and so the Fourier analytic results

$$\|\psi_{h'}\|_{L_2(\mathbb{R})} = \|\bullet\,\psi_h(\bullet)\|_{L_2(\mathbb{R})} \text{ and } \psi_{\psi_h} = 2\pi\, h(-\bullet)$$

can be used. So we have

$$\begin{aligned} +\infty > Const. &\ge \|\psi'_{h_n}\|^2_{L_2(\mathbb{R})} + \|\psi'_{\tilde{h}_n}\|^2_{L_2(\mathbb{R})} \\ &= \frac{1}{2\pi}\|\psi_{\psi'_{h_n}}\|^2_{L_2(\mathbb{R})} + \frac{1}{2\pi}\|\psi_{\psi'_{\tilde{h}_n}}\|^2_{L_2(\mathbb{R})} \\ &= \frac{1}{2\pi}\int |t|^2|\psi_{\psi_{h_n}}(t)|^2 dt + \frac{1}{2\pi}\int |t|^2|\psi_{\psi_{\tilde{h}_n}}(t)|^2 dt \\ &= \int |t|^2 h_n(-t)^2 dt + \int |t|^2 \tilde{h}_n(-t)^2 dt \\ &= \int t^2 h_n(t)^2 dt + \int t^2 \tilde{h}_n(t)^2 dt. \end{aligned} \tag{4.1}$$

This inequality will be applicated later.

The Fourier transform of every density in $\mathcal{F}$ is upper bounded to

$$S(t) = \begin{cases} 1 & , \text{ for } |t| \le T \\ C_2|t|^{-\beta} & , \text{ otherwise.} \end{cases}$$

Since S is square integrable over the whole real line, there is a uniform upper bound for the densities in $\mathcal{F}$ relating to their $L_2(\mathbb{R})$-norm. Let us call this upper bound $C > 0$. Notice that hence one can also postulate $\|\hat{f}_n\|_{L_2(\mathbb{R})} \le C$ without loss of generality.
The MISE of an arbitrary estimator $\hat{f}_n$ can be lower bounded

$$\begin{aligned} &\sup_{g\in\mathcal{G}}\sup_{f\in\mathcal{F}} E_{f*g}\|\hat{f}_n(Y_1,\ldots,Y_n) - f\|^2_{L_2(\mathbb{R})} \\ &\ge \sup_{f\in\mathcal{F}} E_{f*g_N}\|\hat{f}_n(Y_1,\ldots,Y_n) - f\|^2_{L_2(\mathbb{R})} \\ &\ge \tfrac{1}{2}\Big(E_{h_n}\|\hat{f}_n(Y_1,\ldots,Y_n) - f_n\|^2_{L_2(\mathbb{R})} + E_{\tilde{h}_n}\|\hat{f}_n(Y_1,\ldots,Y_n) - \tilde{f}_n\|^2_{L_2(\mathbb{R})}\Big) \\ &\ge \tfrac{1}{2}\Big(E_{h_n}\|\hat{f}_n(Y_1,\ldots,Y_n) - f_n\|^2_{L_2(\mathbb{R})} - E_{\tilde{h}_n}\|\hat{f}_n(Y_1,\ldots,Y_n) - f_n\|^2_{L_2(\mathbb{R})} \\ &+ E_{\tilde{h}_n}\|\hat{f}_n(Y_1,\ldots,Y_n) - f_n\|^2_{L_2(\mathbb{R})} + E_{\tilde{h}_n}\|\hat{f}_n(Y_1,\ldots,Y_n) - \tilde{f}_n\|^2_{L_2(\mathbb{R})}\Big) \\ &\ge \tfrac{1}{2}\Big(E_{\tilde{h}_n}\|\hat{f}_n(Y_1,\ldots,Y_n) - f_n\|^2_{L_2(\mathbb{R})} + E_{\tilde{h}_n}\|\hat{f}_n(Y_1,\ldots,Y_n) - \tilde{f}_n\|^2_{L_2(\mathbb{R})} \\ &-\big|E_{h_n}\|\hat{f}_n(Y_1,\ldots,Y_n) - f_n\|^2_{L_2(\mathbb{R})} - E_{\tilde{h}_n}\|\hat{f}_n(Y_1,\ldots,Y_n) - f_n\|^2_{L_2(\mathbb{R})}\big|\Big) \\ &\ge \tfrac{1}{2}\Big(E_{\tilde{h}_n}\Big(\|\hat{f}_n(Y_1,\ldots,Y_n) - f_n\|^2_{L_2(\mathbb{R})} + \|\hat{f}_n(Y_1,\ldots,Y_n) - \tilde{f}_n\|^2_{L_2(\mathbb{R})}\Big) \\ &-\big|\textstyle\int\cdots\int \|\hat{f}_n(y_1,\ldots,y_n) - f_n\|^2_{L_2(\mathbb{R})}|h_n(y_1)\cdots h_n(y_n) - \tilde{h}_n(y_1)\cdots\tilde{h}_n(y_n)|\big|\Big) \\ &\ge \tfrac{1}{2}\Big(\tfrac{1}{2}\|f_n - \tilde{f}_n\|^2_{L_2(\mathbb{R})} - C^2\cdot n\cdot \textstyle\int |h_n(y) - \tilde{h}_n(y)|dy\Big). \end{aligned}$$

Defining the density

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

with an appropriately selected constant $c > 0$, the inequality sequence above continues

$$\begin{aligned}
&\tfrac{1}{2}\Big(\tfrac{1}{2}\|f_n - \tilde{f}_n\|^2_{L_2(\mathbb{R})} - C^2 \cdot n \cdot \textstyle\int |h_n(y) - \tilde{h}_n(y)| dy\Big) \\
&= \tfrac{1}{2}\Big(\tfrac{1}{2}\|f_n - \tilde{f}_n\|^2_{L_2(\mathbb{R})} - C^2 \cdot n \cdot \textstyle\int \sqrt{\xi(y)} |h_n(y) - \tilde{h}_n(y)| / \sqrt{\xi(y)} dy\Big) \\
&\geq \tfrac{1}{2}\Big(\tfrac{1}{2}\|f_n - \tilde{f}_n\|^2_{L_2(\mathbb{R})} - C^2 \cdot n \cdot \underbrace{(\int \xi(y) dy)^{1/2}}_{=1} \cdot (\textstyle\int |h_n(y) - \tilde{h}_n(y)|^2 / \xi(y) dy)^{1/2}\Big),
\end{aligned}$$

utilizing the Cauchy-Schwarz-inequality. Now one can see that if the condition

$$\frac{n \cdot (\int |h_n(y) - \tilde{h}_n(y)|^2 / \xi(y) dy)^{1/2}}{\|f_n - \tilde{f}_n\|^2_{L_2(\mathbb{R})}} \overset{n\to\infty}{\longrightarrow} 0 \tag{4.2}$$

holds then the MISE of the estimator can be lower bounded by

$$\frac{1}{8}\|f_n - \tilde{f}_n\|^2_{L_2(\mathbb{R})}. \tag{4.3}$$

One has to remember that the sequence $(M_n)_n$ can still be chosen appropriately. First, we regard

$$\begin{aligned}
\|f_n - \tilde{f}_n\|^2_{L_2(\mathbb{R})} &= \tfrac{1}{2\pi} \textstyle\int |\varphi_n(t) - \tilde{\varphi}_n(t)|^2 \, dt \\
&= \tfrac{1}{2\pi} \textstyle\int_{|t| \geq M_n} |\varphi_n(t) - \tilde{\varphi}_n(t)|^2 \, dt \\
&\leq \tfrac{C_2^2}{\pi} \textstyle\int_{|t| \geq M_n} t^{-4} \, dt \\
&\leq Const. \cdot M_n^{-3}.
\end{aligned}$$

On the other hand, one can also derive a lower bound for this term

$$\begin{aligned}
\|f_n - \tilde{f}_n\|^2_{L_2(\mathbb{R})} &= \tfrac{1}{2\pi} \textstyle\int |\varphi_n(t) - \tilde{\varphi}_n(t)|^2 \, dt \\
&\geq \tfrac{1}{2\pi} \textstyle\int_{|t| \geq D \cdot M_n} (C_2 - C_1)^2 t^{-4} \, dt \\
&\geq Const. \cdot M_n^{-3}.
\end{aligned}$$

So we have $\|f_n - \tilde{f}_n\|^2_{L_2(\mathbb{R})} \sim M_n^{-3}$.

Now, we consider

$$\begin{aligned}
\|h_n - \tilde{h}_n\|^2_{L_2(\mathbb{R})} &= \tfrac{1}{2\pi} \textstyle\int |\psi_{g_N}(t)|^2 |\varphi_n(t) - \tilde{\varphi}_n(t)|^2 \, dt \\
&= \tfrac{1}{2\pi} \textstyle\int_{|t| \geq M_n} |\psi_{g_N}(t)|^2 |\varphi_n(t) - \tilde{\varphi}_n(t)|^2 \, dt \\
&\geq \tfrac{C_2^2}{2\pi} \exp\big(-M_n^2\big) \textstyle\int_{|t| \geq M_n} t^{-4} dt \\
&\geq Const. \cdot \exp\big(-M_n^2\big) M_n^{-3}.
\end{aligned}$$

Hence, with $(R_n)_n$ being a sequence yet to be specified

$$
\begin{aligned}
& n \cdot \big(\textstyle\int |h_n(y) - \tilde{h}_n(y)|^2/\xi(y)dy\big)^{1/2} \\
& = n \cdot \Big(\int_{|y|\le R_n} |h_n(y) - \tilde{h}_n(y)|^2/\xi(y)dy + \int_{|y|>R_n} |h_n(y) - \tilde{h}_n(y)|^2/\xi(y)dy\Big)^{1/2} \\
& \le n \cdot \Big(\int_{|y|\le R_n} |h_n(y) - \tilde{h}_n(y)|^2/ \underbrace{\xi(y)}_{\ge cR_n^{-3/2}} dy\Big)^{1/2} \\
& \quad + n \cdot \Big(\int_{|y|>R_n} |h_n(y) - \tilde{h}_n(y)|^2/\xi(y)dy\Big)^{1/2} \\
& \le n \cdot c^{-1/2} R_n^{3/4} \cdot \|h_n - \tilde{h}_n\|_{L_2(\mathbb{R})} + n \cdot \Big(\int_{|y|>R_n} |h_n(y) - \tilde{h}_n(y)|^2/\xi(y)dy\Big)^{1/2} \\
& \le Const. \cdot n \cdot R_n^{3/4} M_n^{-3/2} \exp\big(-\tfrac{1}{2}M_n^2\big) \\
& \quad + 2n \cdot \Big(\int_{|y|>R_n} |h_n(y)|^2/\xi(y)dy + \int_{|y|>R_n} |\tilde{h}_n(y)|^2/\xi(y)dy\Big)^{1/2}.
\end{aligned}
$$

The two summands which remain to be upper bounded can be calculated utilizing (4.1).

$$
\begin{aligned}
& \int_{|y|>R_n} |h_n(y)|^2/\xi(y)dy + \int_{|y|>R_n} |\tilde{h}_n(y)|^2/\xi(y)dy \\
& = \int_{|y|>R_n} |h_n(y)|^2 c^{-1} |y|^{3/2} dy + \int_{|y|>R_n} |\tilde{h}_n(y)|^2 c^{-1} |y|^{3/2} dy \\
& \le Const. \cdot \Big(R_n^{-1/2} \int_{|y|>R_n} |h_n(y)|^2 y^2 dy + R_n^{-1/2} \int_{|y|>R_n} |\tilde{h}_n(y)|^2 y^2 dy \Big) \\
& \le Const. \cdot R_n^{-1/2} \cdot \Big(\textstyle\int h_n(y)^2 y^2 dy + \int \tilde{h}_n(y)^2 y^2 dy \Big) \\
& \le Const. \cdot R_n^{-1/2}
\end{aligned}
$$

Finally, we get

$$
\frac{n \cdot (\int |h_n(y) - \tilde{h}_n(y)|^2/\xi(y)dy)^{1/2}}{\|f_n - \tilde{f}_n\|^2_{L_2(\mathbb{R})}} \le Const. \cdot n \cdot \frac{R_n^{3/4} M_n^{-3/2} \exp\big(-\frac{1}{2}M_n^2\big) + R_n^{-1/4}}{M_n^{-3}}.
$$

If one selects $R_n = n^5$ and $M_n = 4(\ln n)^{1/2}$, this term converges to zero and (4.2) is fulfilled. Hence, by (4.3), one finally receives the asymptotic lower bound

$$
M_n^{-3} \sim (\ln n)^{-3/2}.
$$

■

References:

Butucea, C. and Matias, C. (2003). Minimax estimation of the noise level and of the signal density in a semiparametric convolution model, *Submitted*.

Carroll, R.J. and Hall, P. (1988). Optimal rates of convergence for deconvolving a density, *Journal of the American Statistical Association*, **83**, 1184-1186.

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

Devroye, L. (1989). Consistent Deconvolution in density estimation, *Canadian Journal of Statistics*, **17**, 235-239.

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

Efromovich, S. (1997). Density estimation for the case of supersmooth measurement error, *Journal of the American Statistical Association*, **92**, 526-535.

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

Fan, J. (1992). Deconvolution with supersmooth distributions, *Canadian Journal of Statistics*, **20**, 155-169.

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

Goldenshluger, A. (1998). On pointwise adaptive nonparametric deconvolution, *Bernoulli*, **5**, 907-926.

Hesse, C.H. and Meister, A. (2001). Optimal Iterative Density Deconvolution, *Submitted.*

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, *Canadian Journal of Statistics*, **17**, 427-438.

Meister, A. (2003a). Robustheitseigenschaften von Dekonvolutionsdichteschätzern bezüglich Missspezifikation der Fehlerdichte, *Dissertation, Universität Stuttgart.*

Meister, A. (2003b) On the effect of misspecifying the error density in a deconvolution problem, *Submitted.*

Neumann, M.H. (1997). On the effect of estimating the error density in nonparametric deconvolution, *Journal of Nonparametric Statistics*, **7**, 307-330.

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

Wand, M.P. and Jones, M.C. (1995). *Kernel Smoothing*, Chapman and Hall, London.

Alexander Meister
Pfaffenwaldring 57
70569 Stuttgart
Germany
E-Mail: meistear@mathematik.uni-stuttgart.de

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.