Bernoulli **22**(1), 2016, 143–192
DOI: 10.3150/14-BEJ626

Adaptive quantile estimation in deconvolution with unknown error distribution

ITAI DATTNER[1], MARKUS REIß[2,*] and MATHIAS TRABS[2,**]

[1]*Department of Statistics, University of Haifa, 199 Abba Khoushy Ave, Mount Carmel, Haifa 3498838, Israel. E-mail: idattner@stat.haifa.ac.il*
[2]*Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany. E-mail: *mreiss@math.hu-berlin.de; **trabs@math.hu-berlin.de*

Quantile estimation in deconvolution problems is studied comprehensively. In particular, the more realistic setup of unknown error distributions is covered. Our plug-in method is based on a deconvolution density estimator and is minimax optimal under minimal and natural conditions. This closes an important gap in the literature. Optimal adaptive estimation is obtained by a data-driven bandwidth choice. As a side result, we obtain optimal rates for the plug-in estimation of distribution functions with unknown error distributions. The method is applied to a real data example.

Keywords: adaptive estimation; deconvolution; distribution function; minimax convergence rates; plug-in estimator; quantile function; random Fourier multiplier

1. Introduction

Nonparametric deconvolution models are of high practical importance and lead to challenging questions in statistical methodology. Let $X_1, \dots, X_n$ be independent random variables with a common Lebesgue density $f : \mathbb{R} \to \mathbb{R}$. Suppose that we merely observe the random variables

$$Y_j = X_j + \varepsilon_j, \qquad j = 1, \dots, n,$$

that is the original (X_j) corrupted by i.i.d. error variables ε_j, independent of (X_j) and with Lebesgue density f_ε. For $\tau \in (0, 1)$ the objective is to estimate the τ-quantile q_τ of the population X from the observations $Y_1, \dots, Y_n$. For practitioners estimated quantiles are very relevant, but they depend in a nonlinear way on the underlying density such that their estimation is not always obvious. Abstractly, quantile estimation in deconvolution is an example of nonlinear functional estimation in ill-posed inverse problems.

Two natural strategies may be pursued. Either a distribution function estimator is inverted or an M-estimation paradigm is applied using a density estimator of f. While the first possibility was studied by Hall and Lahiri [12], the purpose of this paper is the analysis of the second in a far more general setting. Assuming that the distribution of the measurement error is completely known, Carroll and Hall [1] have constructed a kernel density estimator based on the empirical characteristic function $\varphi_n(u) := \frac{1}{n}\sum_{j=1}^n \mathrm{e}^{\mathrm{i}uY_j}$, $u \in \mathbb{R}$. In practice, however, the distribution of

1350-7265

the measurement error is usually not known. Instead, we assume that we have at hand a sample from f_ε given by

$$\varepsilon_1^*, \dots, \varepsilon_m^*, \qquad m \in \mathbb{N}.$$

Motivated from applications, we will not assume that the observations (ε_k^*) are independent from (Y_j). In particular, our procedure applies to the experimental setup of repeated measurements, as discussed below.

Let $\mathcal{F}g(u) := \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}ux} g(x)\,\mathrm{d}x$, $u \in \mathbb{R}$, denote the Fourier transform of $g \in L^1(\mathbb{R}) \cup L^2(\mathbb{R})$. Consequently, $\mathcal{F}^{-1}[h(u)](x) = \frac{1}{2\pi}\int \mathrm{e}^{-\mathrm{i}ux} h(u)\,\mathrm{d}u, x \in \mathbb{R}$. Based on the classical kernel estimator, Neumann [21] has proposed the following density estimator of f for the case of unknown error distributions:

$$\widetilde{f}_b(x) := \mathcal{F}^{-1}\left[\frac{\varphi_n(u)\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)} \mathbb{1}_{\{|\varphi_{\varepsilon,m}(u)| \geq m^{-1/2}\}}\right](x), \qquad x \in \mathbb{R},$$

where φ_K is the Fourier transform of a kernel K, $b > 0$ is its bandwidth and the characteristic function of the error distribution φ_ε is estimated by its empirical counterpart $\varphi_{\varepsilon,m}(u) := \frac{1}{m}\sum_{k=1}^m \mathrm{e}^{\mathrm{i}u\varepsilon_k^*}$, $u \in \mathbb{R}$. Obviously, $\widetilde{f}_b$ depends on the sample sizes n and m which are suppressed in the notation. Applying a plug-in approach, our estimator for the quantile q_τ is then given by the minimum-contrast estimator

$$\widetilde{q}_{\tau,b} := \underset{\eta \in [-U_n, U_n]}{\arg\min} \left|\widetilde{M}_b(\eta)\right| \qquad \text{with } \widetilde{M}_b(\eta) = \int_{-\infty}^{\eta} \widetilde{f}_b(x)\,\mathrm{d}x - \tau \tag{1}$$

for some $U_n \to \infty$. We will show as the very first step that $\widetilde{f}_b$ is indeed integrable with overwhelming probability and when not, we define $\widetilde{q}_\tau$ to be the empirical τ-quantile of the observations Y_j's. In this work we pursue the analysis for error distributions whose characteristic function decays polynomially. As shown by Fan [8], these so-called ordinary smooth errors lead to mildly ill-posed estimation problems. They are mathematically more challenging than the so-called super-smooth errors, which we discuss briefly in Section 2.3.

Although the literature on deconvolution problems is extensive and very broad, the problem of adaptive deconvolution with unknown measurement errors was addressed only recently, see Comte and Lacour [3], Johannes and Schwarz [14] and Kappus [16] for adaptive density estimation with unknown error distributions in the model selection framework. Minimax results and other properties for nonadaptive methods are given by Neumann [21,22], Meister [20], Delaigle, Hall and Meister [6], Johannes [13] among others. To the best of our knowledge, the problem of quantile estimation in deconvolution was considered only in Hall and Lahiri [12]. They have constructed a quantile estimator for the case of known error distributions by inverting the distribution function estimator, without proposing an adaptive bandwidth choice. As we shall establish, the error of the quantile estimator (1) is directly related to that of the distribution function estimator (cf. the error representation (4) below). Yet, the general analysis of the latter was not clear before.

Fan [8] has proposed an estimator for the distribution function by integrating the density deconvolution estimator. In order to perform an exact analysis of its variance, a truncation of the integral was required in the estimation procedure. This resulted in a nonoptimal (in the minimax sense) estimation method for the case of ordinary smooth errors and raised the conjecture that ‘plug-in does not work optimally’ for estimation of the distribution function in deconvolution. Trying to circumvent this problem, Hall and Lahiri [12] as well as Dattner, Goldenshluger and Juditsky [4] have constructed a distribution function estimator based on a direct inversion formula. Applying the Fourier multiplier approach by Nickl and Reiß [24], Söhl and Trabs [26] have shown that the integrated density estimator can indeed estimate the distribution function with $\sqrt{n}$-rate under suitable conditions. Since they prove a Donsker theorem, the imposed conditions are restrictive. In particular, a global Sobolev regularity of f is assumed there which is not natural for pointwise loss. So even with a known error distribution, it remained an open and intriguing question whether the canonical plug-in estimator for distribution or quantile function estimation yields asymptotically optimal results under natural conditions.

In Section 2, we settle this question in the positive under local Hölder regularity of f by combining an exact analysis like in Dattner, Goldenshluger and Juditsky [4] together with abstract Fourier multiplier theory from Söhl and Trabs [26]. Moreover, we show that the optimal rates continue to hold if the error distribution is unknown and has to be estimated, which is mathematically nontrivial. Since the deconvolution operator $\mathcal{F}^{-1}[1/\varphi_\varepsilon]$ is not observable, we have to study the estimated counterpart $\mathcal{F}^{-1}[\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\mathbb{1}_{\{|\varphi_{\varepsilon,m}(u)|\geq m^{-1/2}\}}]$. As a random Fourier multiplier, it preserves the mapping properties of the deterministic $\mathcal{F}^{-1}[1/\varphi_\varepsilon]$, but its operator norm turns out to be (slightly) larger.

A lower bound result establishes that the rates under a local Hölder condition are indeed minimax optimal. Surprisingly, the dependence of the minimax rate on the error sample size m is completely different from the case of global Sobolev restrictions like in Neumann [21]. The proof enlightens this interplay between the decay of one characteristic function and estimation error in the other sample for both, the (Y_j) and the (ε_j).

An adaptive (data-driven) bandwidth choice is developed in Section 3. To this end, a variant of Lepski’s method is applied, but because of the unknown and possibly dependent error distribution a much more refined analysis is needed to establish that the resulting adaptive quantile estimator is (up to log factors) still rate optimal.

In Section 4, we implement our estimation procedure and present simulation results which show a good performance of the estimator. In a real data example, we consider multiple blood pressure measurement data from different patients. Here, a measurement error is clearly present, but of unknown distribution and we have to estimate it by taking patient-wise differences. The completely data-driven method yields reasonable quantile estimates which differ from the sample quantiles of the directly measured (Y_j). All proofs are postponed to Section 5.

2. Convergence rates

2.1. Setting and upper bounds

Let us introduce some notation. Denoting $\langle\alpha\rangle$ as the largest integer which is strictly smaller than $\alpha > 0$, we define for some function g and any possibly unbounded interval $I \subseteq \mathbb{R}$ the Hölder norm

$$\|g\|_{C^\alpha(I)} := \sum_{k=0}^{\langle\alpha\rangle} \|g^{(k)}\|_{L^\infty(I)} + \sup_{x,y\in I\,:\,x\neq y} \frac{|g^{\langle\alpha\rangle}(x) - g^{\langle\alpha\rangle}(y)|}{|x-y|^{\alpha-\langle\alpha\rangle}}.$$

Let $C^0(I)$ denote the space of all continuous and bounded functions on the interval I and

$$C^s(\mathbb{R}) := \bigcup_{R>0} C^s(\mathbb{R}, R) \qquad \text{with } C^\alpha(I, R) := \{g \in C^0(I) \mid \|g\|_{C^\alpha(I)} \le R\}, R > 0.$$

In the sequel, we use the Landau notation $\mathcal{O}$ and $\mathcal{O}_P$. For two sequences $A_n(\vartheta), B_n(\vartheta)$ depending on a parameter ϑ, $A_n(\vartheta) = \mathcal{O}_P(B_n(\vartheta))$ holds uniformly over a parameter set $\vartheta \in \Theta$ if there is for all $c > 0$ some $C > 0$ such that $\sup_{\vartheta\in\Theta} P_\vartheta(A_n(\vartheta) > CB_n(\vartheta)) < c$. If $A_n(\vartheta)/B_n(\vartheta)$ converges in probability to zero, we write $A_n(\vartheta) = \mathrm{o}_P(B_n(\vartheta))$.

Assumption A. *Let the kernel function $K \in L^1(\mathbb{R})$ with Fourier transform $\varphi_K := \mathcal{F}K$ satisfy*

(i) $\operatorname{supp}\varphi_K \subseteq [-1, 1]$ *and*
(ii) *K has order $\ell \in \mathbb{N}$, i.e., for $k = 0, \dots, \ell$*

$$\int_{\mathbb{R}} |K(x)||x|^{\ell+1}\,\mathrm{d}x < \infty \quad \textit{and} \quad \int_{\mathbb{R}} x^k K(x)\,\mathrm{d}x = \begin{cases} 1, & \textit{if } k = 0, \\ 0, & \textit{otherwise.} \end{cases}$$

By construction the quantile estimator, $\widetilde{q}_{\tau,b}$ is the approximated solution of the estimating equation

$$0 = \widetilde{M}_b(\eta) = \int_{-\infty}^{\eta} \widetilde{f}_b(x)\,\mathrm{d}x - \tau. \tag{2}$$

If a solution exists, it does not have to be unique since $\widetilde{f}_b$ is not necessarily nonnegative. Nevertheless, any choice converges to the true quantile, assuming the latter is unique. Before, integrability of $\widetilde{f}_b$ was an open problem, which we shall settle now.

Lemma 2.1. *Grant Assumption A with $\ell = 0$. On the event*

$$B_\varepsilon(b) := \Big\{ \inf_{u\in[-1/b,1/b]} |\varphi_{\varepsilon,m}(u)| \ge m^{-1/2} |\log b|^{3/2} \Big\} \tag{3}$$

we have $\widetilde{f}_b \in L^1(\mathbb{R})$ and estimating equation (2) has a solution.

Therefore, a truncation of the integral as used by Fan [8] is not necessary, implying that no tail condition on f is required. Although $\|\widetilde{f}_b\|_{L^1}$ is finite, it depends on the observations as well as through b on n, m. To quantify the behavior of $\widetilde{f}_b$ more precisely, our analysis relies on the following much stronger result.

Lemma 2.2. *Grant Assumption A with $\ell = 0$. For some $\beta, R > 0$ suppose $\mathbb{E}[(\varepsilon_k^*)^4] \le R$ and*

$$|\varphi_\varepsilon(u)|^{-1} \le R(1 + |u|)^\beta \quad \textit{and} \quad |\varphi_\varepsilon'(u)| \le R(1 + |u|)^{-\beta-1}$$

as well as $mb^{2\beta+1} \to \infty$. Then there exists a finite random variable $\mathcal{E}_b$ which is $\mathcal{O}_P(1 \vee \frac{1}{m^{1/2}b^{\beta+1}})$ with the constant depending only on β and R, such that for any $s > \beta^+ > \beta$ on the event $B_\varepsilon(b)$ from (3)

$$\left\| \mathcal{F}^{-1}\left[\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\right] * \psi \right\|_{C^{s-\beta^+}(\mathbb{R})} \le \mathcal{E}_b \|\psi\|_{C^s(\mathbb{R})} \qquad \textit{for all } \psi \in C^s(\mathbb{R}).$$

The deterministic counterpart of this lemma was proved by Söhl and Trabs [26]. Here, we show that the random Fourier multiplication operator $C^s(\mathbb{R}) \ni \psi \mapsto \mathcal{F}^{-1}[\frac{\varphi_K(bu)\mathcal{F}\psi(u)}{\varphi_{\varepsilon,m}(u)}] \in C^{s-\beta^+}(\mathbb{R})$ has a norm bound $\mathcal{O}_P(1 \vee \frac{1}{m^{1/2}b^{\beta+1}})$ on the event $B_\varepsilon(b)$. The condition on the derivative φ_ε' is natural in the context of Fourier multipliers and is usually satisfied for distributions with polynomial decaying characteristic functions, for example, Gamma distributions with shape parameter $\beta > 0$ satisfy it.

Remark 2.3. Depending only on the observations, condition (3) can be verified by the practitioner for a given bandwidth b. Under the assumptions of Lemma 2.2 Talagrand's inequality yields $P(B_\varepsilon(b)) \ge 1 - 2\mathrm{e}^{-mb^{2\beta+1}}$ (cf. Lemma 5.1 and (53) below). Therefore, with overwhelming probability $B_\varepsilon(b)$ holds true and the estimating equation (2) is rigorously defined.

Before we start with the error analysis, let us describe the class of densities we are interested in. Let $\mathcal{Q}(R)$ denote the set of all probability densities on $\mathbb{R}$ which are uniformly bounded by $R > 0$. Following the minimax paradigm, we consider for $R, r, \zeta, U > 0$ and the smoothness index $\alpha > 0$ the classes

$$\mathcal{C}^\alpha(R, r, \zeta) := \bigcup_{U \in \mathbb{N}} \mathcal{C}^\alpha(R, r, \zeta, U) \quad \text{and}$$

$$\mathcal{C}^\alpha(R, r, \zeta, U) := \big\{ f \in \mathcal{Q}(R) | f \text{ has a } \tau\text{-quantile } q_\tau \in [-U, U] \text{ such that } f \in C^\alpha\big([q_\tau - \zeta, q_\tau + \zeta], R\big) \text{ and } f(q_\tau) \ge r \big\}.$$

In contrast to Dattner, Goldenshluger and Juditsky [4], the smoothness is measured locally in a Hölder scale and not globally by decay conditions of the Fourier transform of f. The former is more natural since both, the distribution function and the quantile function are estimated pointwise. Note that the quantile q_τ is unique given the assumption $f(q_\tau) > 0$. Recalling that we write $\varphi_\varepsilon := \mathcal{F} f_\varepsilon$, the conditions in Lemma 2.2 motivate the definition of the class of error densities

$$\mathcal{D}^\beta(R, \gamma) := \bigg\{ f_\varepsilon \in \mathcal{Q}(\infty) \bigg| \frac{1}{R}\big(1 + |u|\big)^{-\beta} \le \big|\mathcal{F} f_\varepsilon(u)\big| \le R\big(1 + |u|\big)^{-\beta}, \\ \big|(\mathcal{F} f_\varepsilon)'(u)\big| \le R\big(1 + |u|\big)^{-1-\beta}, \ \big\|x^\gamma f_\varepsilon(x)\big\|_{L^1} \le R \bigg\}$$

for some moment $\gamma \ge 0$ and we use the same constant R as above for convenience.

Remark 2.4. The upper and lower bounds for $|\varphi_\varepsilon(u)|$ in $\mathcal{D}^\beta(R,\gamma)$ are standard assumptions in deconvolution and are used for deriving lower bounds for the estimation problem as well as upper bounds for the risk of the estimators. Specifically, these bounds correspond to ordinary smooth error distributions (Fan [8]), cf. Section 2.3 below for the super-smooth case.

Applying the plug-in approach, we need to integrate the density estimator over an unbounded interval. As mentioned above, additional assumptions are necessary to control $\|\widetilde{f}_b\|_{L^1}$. We apply Lemma 2.2 assuming $\gamma \geq 4$, that is $\mathbb{E}[(\varepsilon_1^*)^4] < \infty$, and a polynomial decay of $|\varphi_\varepsilon'|$. The latter is a natural Mihlin-type condition in the context of Fourier multipliers. Note that φ_ε' exists if f_ε, the distribution of the measurement errors, has a first moment. In view of the analysis by Neumann and Reiß [23], the moment assumption in particular implies uniform convergence of $\varphi_{\varepsilon,m}$.

To control the estimation error of $\widetilde{q}_{\tau,b}$, we follow the Z-estimator approach (cf. van der Vaart [31]). Let $M(\eta)$ be the deterministic counterpart of $\widetilde{M}_b(\eta)$ defined in (1). The quantities $\widetilde{q}_{\tau,b}$ and q_τ are given by the (approximated) zeros of $\widetilde{M}_b$ and M, respectively. From the Taylor expansion $0 \approx \widetilde{M}_b(\widetilde{q}_{\tau,b}) = \widetilde{M}_b(q_\tau) + (\widetilde{q}_{\tau,b} - q_\tau)\widetilde{M}_b'(q_\tau^*)$ for some intermediate point q_τ^* between q_τ and $\widetilde{q}_{\tau,b}$, we obtain

$$\widetilde{q}_{\tau,b} - q_\tau \approx -\frac{\int_{-\infty}^{q_\tau}(\widetilde{f}_b(x) - f(x))\,\mathrm{d}x}{\widetilde{f}_b(q_\tau^*)}. \tag{4}$$

The following two propositions deal separately with the numerator and the denominator in this representation. The results are intrinsic to our analysis, but may also be of interest on their own. The first proposition deals with the numerator in (4) and establishes minimax rates of convergence for estimation of the distribution function with unknown error distributions. Note that the quotient in (4) might explode if $\widetilde{f}_b(q_\tau^*)$ becomes very small for large stochastic error. Excluding this event which has vanishing probability, we establish convergence rates as $\mathcal{O}_P$-results.

Proposition 2.5. *Suppose that Assumption A holds with $\ell = \langle\alpha\rangle + 1$ and let $b_{n,m}^* = (n \wedge m)^{-1/(2\alpha+2(\beta\vee 1/2)+1)}$. Then for any $\alpha \geq 1/2$, $\beta, R, r, \zeta > 0$ and $\gamma \geq 4$ we have uniformly over $f \in \mathcal{C}^\alpha(R,r,\zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R,\gamma)$ as $n \wedge m \to \infty$,*

$$\left|\int_{-\infty}^{q_\tau}\big(\widetilde{f}_{b_{n,m}^*}(x) - f(x)\big)\,\mathrm{d}x\right| = \mathcal{O}_P\big(\psi_{n\wedge m}(\alpha,\beta)\big),$$

where for $k \geq 1$

$$\psi_k(\alpha,\beta) := \begin{cases} k^{-1/2}, & \text{for } \beta \in (0,1/2), \\ (\log k/k)^{1/2}, & \text{for } \beta = 1/2, \\ k^{-(\alpha+1)/(2\alpha+2\beta+1)}, & \text{for } \beta > 1/2. \end{cases} \tag{5}$$

Since the techniques to obtain Proposition 2.5 differ significantly from previous results for deconvolution with unknown error distribution, let us briefly sketch the proof: we apply a smooth truncation function a_s to decompose the error into

$$\begin{aligned}
&\int_{-\infty}^{q_\tau} \big(\widetilde{f}_b(x) - f(x)\big)\,\mathrm{d}x \\
&\quad = \underbrace{\int_{-\infty}^{q_\tau} \big(K_b * f(x) - f(x)\big)\,\mathrm{d}x}_{\text{deterministic error}} + \underbrace{\int_{-\infty}^{q_\tau} a_s(x+q_\tau)\big(\widetilde{f}_b(x) - K_b * f(x)\big)\,\mathrm{d}x}_{\text{singular part of stochastic error}} \qquad (6) \\
&\qquad + \underbrace{\int_{-\infty}^{q_\tau} \big(1 - a_s(x+q_\tau)\big)\big(\widetilde{f}_b(x) - K_b * f(x)\big)\,\mathrm{d}x}_{\text{continuous part of stochastic error}}
\end{aligned}$$

with the usual notation $K_b(\cdot) = b^{-1}K(\cdot/b)$. The function a_s can be chosen such that it has compact support and satisfies $(\mathbb{1}_{(-\infty,0]} - a_s) \in C^\infty(\mathbb{R})$. Similar to the classical bias-variance trade-off, the deterministic error and singular part of the stochastic error will determine the rate. The continuous part, however, corresponds to the estimation error of a smooth (but not integrable) functional of the density. If the error distribution were known, it would be of order $n^{-1/2}$. For unknown errors we use Lemma 2.2, where our estimate of the operator norm of the random Fourier multiplier $\mathcal{F}^{-1}[\varphi_K(bu)/\varphi_{\varepsilon,m}(u)\mathbb{1}_{\{|\varphi_{\varepsilon,m}(u)| \geq m^{-1/2}\}}]$ is of order $\mathcal{O}_P(1 \vee (m^{-1/2}b^{-\beta-1}))$. This might be larger than the operator norm of the unknown deconvolution operator $\mathcal{F}^{-1}[1/\varphi_\varepsilon(u)]$ which is uniformly bounded. Yet, for $\alpha \geq 1/2$ the additional error that appears in the continuous part of stochastic error in (6) is negligible.

Next, we like to understand the denominator of (4). Lounici and Nickl [18] have proved uniform risk bounds for the deconvolution wavelet estimator on the whole real line for a known error distribution. On a bounded interval, which is sufficient for our purpose, uniform convergence of the deconvolution estimator $\widetilde{f}_b$ can be proved more elementarily. With $b_n = (\log n/n)^{1/(2\alpha+2\beta+1)}$ the following proposition yields the minimax rate $(\log n/n)^{\alpha/(2\alpha+2\beta+1)}$ in L^∞-loss (at least if $\frac{n}{\log n} \leq m$).

Proposition 2.6. *Grant Assumption A with $\ell = \langle\alpha\rangle$. For any $\alpha, \beta, R, r, \zeta > 0$ and $\gamma \geq 0$ we have uniformly over $f \in \mathcal{C}^\alpha(R, r, \zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$ as $n \wedge m \to \infty$,*

$$\sup_{x \in (-\zeta,\zeta)} \big|\widetilde{f}_b(x+q_\tau) - f(x+q_\tau)\big| = \mathcal{O}_P\left(b^\alpha + \left(\frac{\log n}{n} \vee \frac{1}{m}\right)^{1/2} b^{-\beta-1/2}\right).$$

In particular, if $b = b_{n,m} \to 0$ and $(\frac{n}{\log n} \wedge m)b_{n,m}^{2\beta+1} \to \infty$ as $n \wedge m \to \infty$, $\widetilde{f}_{b_{n,m}}$ is a uniformly consistent estimator.

The two propositions above are the building blocks for the first main result of this paper announced in the following theorem. The constant preceding the rate depends only on the class parameters $\alpha, \beta, \gamma, R, r, \zeta$. The location parameter U_n can grow logarithmically to infinity as $n \to \infty$.

Theorem 2.7. *Let $\alpha \geq 1/2$, $\beta, R, r, \zeta > 0$ and $\gamma \geq 4$ and grant Assumption A with $\ell = \langle\alpha\rangle + 1$. Let $\widetilde{q}_{\tau,b^*_{n,m}}$ be the quantile estimator defined in (1) associated with $b^*_{n,m} = (n \wedge m)^{-1/(2\alpha+2(\beta\vee 1/2)+1)}$ and with $U_n \to \infty, U_n = \mathcal{O}(\log n)$. Then we have uniformly over $f \in \mathcal{C}^\alpha(R, r, \zeta, U_n)$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$ as $n \wedge m \to \infty$,*

$$|\widetilde{q}_{\tau,b^*_{n,m}} - q_\tau| = \mathcal{O}_P\big(\psi_{n\wedge m}(\alpha, \beta)\big),$$

where $\psi_\cdot(\alpha, \beta)$ is given in (5).

Using the methods of the proof of Theorem 2.7 and an additional application of Bernstein's concentration inequality, convergence rates for the uniform loss can be obtained, assuming regularity in a neighborhood of some interval of quantiles. For $0 < \tau_1 < \tau_2 < 1$ and $\alpha, R, r, \zeta, U_n > 0$, define

$$\begin{aligned}
&\mathcal{C}^\alpha_\infty(\tau_1, \tau_2, R, r, \zeta, U_n) \\
&\quad := \Big\{ f \in \mathcal{Q}(R) \Big| \text{for all } \tau \in (\tau_1, \tau_2)\text{: } f \text{ has a } \tau\text{-quantile } q_\tau \in [-U_n, U_n] \text{ and} \\
&\qquad f \in C^\alpha\big([q_{\tau_1} - \zeta, q_{\tau_2} + \zeta], R\big), \inf_{\tau\in(\tau_1,\tau_2)} f(q_\tau) \geq r \Big\}.
\end{aligned}$$

Theorem 2.8. *Let $\alpha \geq 1/2$, $\beta, R, r, \zeta > 0$ and $\gamma \geq 4$ and grant Assumption A with $\ell = \langle\alpha\rangle + 1$. For $0 < \tau_1 < \tau_2 < 1$ and $\tau \in (\tau_1, \tau_2)$ let $\widetilde{q}_{\tau,b^*_{n,m}}$ be the quantile estimator defined in (1) associated with $b^*_{n,m} = (\frac{\log n}{n} \vee \frac{1}{m})^{1/(2\alpha+2(\beta\vee 1/2)+1)}$ and with $U_n \to \infty, U_n = \mathcal{O}(\log n)$. Then we have uniformly over $f \in \mathcal{C}^\alpha_\infty(\tau_1, \tau_2, R, r, \zeta, U_n)$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$ as $n \wedge m \to \infty$,*

$$\sup_{\tau\in(\tau_1,\tau_2)} |\widetilde{q}_{\tau,b^*_{n,m}} - q_\tau| = \mathcal{O}_P\big(\psi_{(n/\log n)\wedge m}(\alpha, \beta)\big),$$

where $\psi_\cdot(\alpha, \beta)$ is given in (5).

We finish this subsection by providing the minimax rates for estimating the distribution function and the quantiles for the case of known error distributions, restricting to pointwise loss. As above, the estimators are given by plugging in the classical density estimator

$$\widehat{f}_b(x) := \mathcal{F}^{-1}\bigg[\frac{\varphi_n(u)\varphi_K(bu)}{\varphi_\varepsilon(u)}\bigg](x), \qquad x \in \mathbb{R}. \tag{7}$$

Corollary 2.9. *Let $\alpha, \beta, R, r, \zeta > 0$ and $\gamma \geq 0$ and suppose that the error distribution is known and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$. Let Assumption A hold with $\ell = \langle\alpha\rangle + 1$. Let $\widehat{q}_{\tau,b}$ be the quantile estimator based on the density deconvolution estimator (7) associated with $b^*_n = n^{-1/(2\alpha+2(\beta\vee 1/2)+1)}$ and $U_n \to \infty, U_n = \mathcal{O}(\log n)$. Then we obtain uniformly over $f \in \mathcal{C}^\alpha(R, r, \zeta, U_n)$ as $n \to \infty$,*

$$\bigg|\int_{-\infty}^{q_\tau} \big(\widehat{f}_{b^*_n}(x) - f(x)\big)\,\mathrm{d}x\bigg| = \mathcal{O}_P\big(\psi_n(\alpha, \beta)\big),$$

$$|\widehat{q}_{\tau,b^*_n} - q_\tau| = \mathcal{O}_P\big(\psi_n(\alpha, \beta)\big),$$

where $\psi_{\cdot}(\alpha, \beta)$ is given (5).

Here, we do not estimate the deconvolution operator and thus there is no additional error in terms of m. Consequently, we do not need a moment assumption on the error distribution and the convergence rates hold true for all $\alpha > 0$.

2.2. Lower bounds

In view of the lower bounds stated by Fan [8], in case $n \le m$ the rates in Proposition 2.5 are optimal. Using the error representation (4), the result for distribution function estimation carries over to quantile estimation. Therefore, we focus on the case $m < n$. To provide a clear proof of the lower bound, we allow for a more general class of distributions of X_j, assuming only local assumptions. Using point measures, the estimation error of φ_ε does not profit from the decay of the characteristic function of X_j. One could also consider the case of bounded densities f and choose alternatives in the proof whose Fourier transforms decay arbitrarily slowly, but this would require far more technical arguments.

We define for $\alpha, R, r, \zeta > 0$ and some interval $I \subseteq \mathbb{R}$

$$\widetilde{\mathcal{C}}^{\alpha+1}(R, r, I) := \Big\{ F \text{ c.d.f.} \Big| F \text{ has on } I \text{ a Lebesgue density } f \in C^\alpha(I, R) \text{ and } \inf_{x \in I} f(x) \ge r \Big\},$$

$$\widetilde{\mathcal{C}}^{\alpha+1}(R, r, \zeta) := \big\{ F \text{ c.d.f.} | F \text{ has a } \tau\text{-quantile } q_\tau \in \mathbb{R} \text{ and } F \in \widetilde{\mathcal{C}}^{\alpha+1}\big(R, r, [q_\tau - \zeta, q_\tau + \zeta]\big)\big\}.$$

Theorem 2.10. *Suppose that $Y_1, \dots, Y_n$ and $\varepsilon_1^*, \dots, \varepsilon_m^*$ are independent. Let $q \in \mathbb{R}$ and $\alpha, \beta, R, r, \zeta > 0, \gamma \ge 0$. Then for any $C > 0$ there is some $\delta > 0$ such that*

$$\inf_{\bar{F}_{n,m}} \sup_{F \in \widetilde{\mathcal{C}}^{\alpha+1}(R,r,[q-\zeta,q+\zeta])} \sup_{f_\varepsilon \in \mathcal{D}^\beta(R,\gamma)} P\Big(\big|\bar{F}_{n,m}(q) - F(q)\big| > C(n \wedge m)^{-(\alpha+1)/(2\alpha+(2\beta)\vee 1+1)}\Big) \ge \delta,$$

$$\inf_{\bar{q}_{\tau,n,m}} \sup_{F \in \widetilde{\mathcal{C}}^{\alpha+1}(R,r,\zeta)} \sup_{f_\varepsilon \in \mathcal{D}^\beta(R,\gamma)} P\Big(\big|\bar{q}_{\tau,n,m} - q_\tau\big| > C(n \wedge m)^{-(\alpha+1)/(2\alpha+(2\beta)\vee 1+1)}\Big) \ge \delta,$$

where the infima are taken over all estimators $\bar{F}_{n,m}$ and $\bar{q}_{\tau,n,m}$, respectively.

This lower bound implies that the rates in Proposition 2.5 and Theorem 2.7 are minimax optimal, except for the case $\beta = 1/2$ where they deviate by a logarithmic factor.

2.3. Discussion and extension

The previous results show that estimating the distribution function by integrating a density deconvolution estimator is a minimax optimal procedure and under the local Hölder condition the rates are determined by $n \wedge m$. In that point our results differ completely from previous studies. Assuming α-Sobolev regularity of f, the RMSE of the kernel density estimator by Neumann

[21] is of order $\mathcal{O}(n^{-\alpha/(2\alpha+\beta+1)} + m^{-((\alpha/\beta)\wedge 1)})$. Since the error in estimating φ_ε is reduced by the decay of the characteristic function φ of X_j, the risk is of much smaller order in m. Assuming local regularity on f only, $\mathcal{F}f$ can decay arbitrarily slowly such that this reduction effect may not occur. Note that assuming global Sobolev regularity would improve also the convergence rate of the plug-in estimator.

Interestingly, the dependence on n and m is not completely symmetric. As an intrinsic property of the uniform loss, the convergence rates are typically by a logarithmic factor slower than for pointwise loss. Yet, in Proposition 2.6 and Theorem 2.8 this payment for uniform convergence affects only the estimation of φ and thus the rate is determined by $\frac{\log n}{n} \vee \frac{1}{m}$.

Although the focus of this paper is on ordinary smooth error distributions, a generalization to supersmooth errors is worth mentioning. Let us sketch this case of exponentially decaying φ_ε. Supposing $\mathbb{E}[|\varepsilon_k^*|^4] < \infty$ and $|\varphi_\varepsilon(u)|^{-1} = \mathcal{O}(\mathrm{e}^{\gamma_0|u|^\beta})$ as well as $|\varphi_\varepsilon'(u)| = \mathcal{O}(\mathrm{e}^{-\gamma_1|u|^\beta})$, $u \in \mathbb{R}$, for some $\beta > 0$ and $\gamma_0 \geq \gamma_1 > 0$, we obtain analogously to Lemma 2.2 for sufficiently small $c, \gamma > 0$ and for the bandwidth $b_m^* = c(\log m)^{-1/\beta}$

$$\left\| \mathcal{F}^{-1}\left[\frac{\varphi_K(b_m^* u)}{\varphi_{\varepsilon,m}(u)}\right] * \psi \right\|_{C^s(\mathbb{R})} \mathbb{1}_{B_\varepsilon(b_m^*)} \leq \mathcal{E}_{b_m^*} \|\psi\|_{C^s(\mathbb{R})} \qquad \text{where } \mathcal{E}_b = \mathcal{O}_P\bigl(1 \vee \mathrm{e}^{\gamma b^{-\beta}}\bigr)$$

for any $s \geq 0$ and for any $\psi \in C^s(\mathbb{R})$. In other words, $\varphi_K(bu)/\varphi_{\varepsilon,m}(u)$ is a random Fourier multiplier on Hölder spaces with exponentially increasing operator norm on the event $B_\varepsilon(b)$. Following the lines of the proof of Proposition 2.5, one sees that the singular as well as the continuous part of the stochastic error in (6) are of the order $\mathcal{O}_P((n \wedge m)^{-1/2}\mathrm{e}^{\gamma b^{-\beta}})$. Combined with the estimate for the deterministic error, the choice $b_{n,m}^* = c(\log(n \wedge m))^{-1/\beta}$ yields for $f \in \mathcal{C}^\alpha(R, r, \zeta)$

$$\left| \int_{-\infty}^{q_\tau} \bigl(\widetilde{f}_{b_n^*}(x) - f(x)\bigr)\,\mathrm{d}x \right| = \mathcal{O}_P\bigl(\bigl(\log(n \wedge m)\bigr)^{-(\alpha+1)/\beta}\bigr).$$

Note that for $n \leq m$ this is the minimax rate for distribution function estimation as given in Fan [8]. Therefore, also for supersmooth error distributions the integral domain does not need to be truncated to estimate the distribution function via the plug-in approach.

3. Adaptive estimation

The choice of the bandwidth b is crucial in applications. Therefore, we develop a fully data-driven procedure to determine a good bandwidth. We follow the approach initiated by Lepskiĭ [17]. More precisely, we use the version proposed in Goldenshluger and Nemirovski [11]. For simplicity, we suppose $n = m$ and focus on the pointwise loss in this section.

Let us consider the family of estimators $\{\widetilde{q}_{\tau,b}, b \in \mathcal{B}_n\}$ where $\widetilde{q}_{\tau,b}$ is defined in (1) and $\mathcal{B}_n$ is a finite set of bandwidths. In view of the error representation (4), it is important that $\widetilde{f}_b(\widetilde{q}_{\tau,b})$ is a consistent estimator of $f(q_\tau)$ for all $b \in \mathcal{B}_n$. Therefore, conditions on the bandwidth as in Proposition 2.6 are necessary for the entire set $\mathcal{B}_n$. These depend on the true but unknown degree of ill-posedness β and on α. We keep to the assumption $\alpha > 1/2$ such that the additional error due to bounding the random Fourier multiplier is negligible. Note that the lower bound for the bandwidth is not determined by the variance of the quantile estimator itself but by the variance of the density estimator and the minimal smoothing which results from $\alpha > 1/2$.

Inspired by Comte and Lacour [3], we propose the following construction of a feasible set $\mathcal{B}_n$: for some $L > 1$ define

$$b_{n,j} := n^{-1} L^j \qquad \text{for } j = 0, \dots, N_n \text{ where } N_n \in \mathbb{N} \text{ satisfies } n^{-1} L^{N_n} \sim (\log n)^{-3}.$$

Choosing

$$\widetilde{j}_n := N_n \wedge \min\left\{ j = 0, \dots, N_n - 1 \colon \frac{1}{2} \le \left(\frac{\log n}{n}\right)^{1/2} \int_{-1/b_{n,j}}^{1/b_{n,j}} \frac{\mathbb{1}_{\{|\varphi_{\varepsilon,m}(u)| \ge m^{-1/2}\}}}{|\varphi_{\varepsilon,m}(u)|} \,\mathrm{d}u \le 1 \right\}, \tag{8}$$

the bandwidth set is given by

$$\mathcal{B}_n := \{b_{n,\widetilde{j}_n}, \dots, b_{n,N_n}\}. \tag{9}$$

Note that by construction $\mathcal{B}_n$ is nonempty and it consists of a monotone increasing sequence of bandwidths such that $b_{n,j+1}/b_{n,j}$ is uniformly bounded in $j = \widetilde{j}_n, \dots, N_n$ and $n \ge 1$. Also, for $n \to \infty$ we have $N_n \lesssim \log n$ and $(\log n)^2 b_{n,N_n} \to 0$. The following lemma establishes two additional properties. The latter one ensures that for any $b \in \mathcal{B}_n$ our estimators are consistent.

Lemma 3.1. *Let (Y_j) and (ε_k^*) be distributed according to $f \in \mathcal{C}^\alpha(R, r, \zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$ with $\alpha \ge 1/2, \beta > 0$. Then with probability converging to one, $\widetilde{j}_n < N_n$ and the optimal bandwidth $b_n^* = n^{-1/(2\alpha+2(\beta\vee 1/2)+1)}$ is contained in the interval $[b_{n,\widetilde{j}_n}, b_{n,N_n}]$ as well as $n b_{n,\widetilde{j}_n}^{2\beta+2} \to \infty$.*

Given the bandwidth set, the adaptive estimator is obtained by selection from the family of estimators $\{\widetilde{q}_{\tau,b}, b \in \mathcal{B}_n\}$. As proposed by Lepskiĭ [17] the adaptive choice should mimic the trade-off between deterministic error and stochastic error. The adaptive choice will be given by the largest bandwidth such that the intersection of all confidence sets, which corresponds to smaller bandwidths, is nonempty. As discussed above, it is sufficient to consider the singular part of the stochastic error in (6) only. To estimate the variance of $\widetilde{q}_{\tau,b}$ corresponding to the latter, we define for some $\delta > 0$

$$\widetilde{\Sigma}_b := \frac{(2\sqrt{2} + \delta)\sqrt{\log\log n}\, \max_{\mu \ge b} \widetilde{\sigma}_{\mu,X} + (\delta \log n)^3 \max_{\mu \ge b} \widetilde{\sigma}_{\mu,\varepsilon} + (1+\delta)|\widetilde{M}_b(\widetilde{q}_{\tau,b})|}{|\widetilde{f}_b(\widetilde{q}_{\tau,b})|}, \tag{10}$$

with the truncation function a_s from decomposition (6) and

$$\widetilde{\sigma}_{b,X}^2 = \frac{1}{n^2} \sum_{j=1}^n \left(\int_{-\infty}^0 a_s(x) \mathcal{F}^{-1}\left[\frac{\varphi_K(bu)\mathrm{e}^{\mathrm{i}uY_j}}{\varphi_{\varepsilon,m}(u)}\right](x + \widetilde{q}_{\tau,b}) \,\mathrm{d}x \right)^2 \quad \text{and} \tag{11}$$

$$\widetilde{\sigma}_{b,\varepsilon}^2 = \frac{1}{4\pi^2 m} \int_{-1/b}^{1/b} |\varphi_K(bu)| \left| \frac{\varphi_n(u)}{\varphi_{\varepsilon,m}(u)} \right|^2 \mathrm{d}u \int_{-1/b}^{1/b} |\varphi_K(bu)| \left| \frac{\mathcal{F}a_s(u)}{\varphi_{\varepsilon,m}(u)} \right|^2 \mathrm{d}u. \tag{12}$$

The parameter δ has minor influence and should be chosen close to zero. Note that we apply a monotonization in the numerator of $\widetilde{\Sigma}_b$ by taking maxima of $\widetilde{\sigma}_{\mu,X}$ and $\widetilde{\sigma}_{\mu,\varepsilon}$, respectively. The correction term $|\widetilde{M}_b(\widetilde{q}_{\tau,b})|$ appears only if $\widetilde{q}_{\tau,b}$ is not the exact solution of the estimating equation (2). Define for any $b \in \mathcal{B}_n$

$$\mathcal{U}_b := [\widetilde{q}_{\tau,b} - \widetilde{\Sigma}_b, \widetilde{q}_{\tau,b} + \widetilde{\Sigma}_b].$$

The adaptive estimator is given by

$$\widetilde{q}_\tau := \widetilde{q}_{\tau,\widetilde{b}_n^*} \qquad \text{with } \widetilde{b}_n^* := \max\left\{ b \in \mathcal{B}_n \,\middle|\, \bigcap_{\mu \le b, \mu \in \mathcal{B}_n} \mathcal{U}_\mu \neq \varnothing \right\}. \tag{13}$$

Note that $\widetilde{b}_n^*$ is well defined since the intersection in (13) is nonempty for $b = b_{n,1}$. The following theorem shows that this estimator achieves the minimax rate up to a logarithmic factor. The proof relies on a comparison with an oracle-type choice of the bandwidth. All ingredients, though, have to be estimated and the dependence between Y_j and ε_k^* requires special attention.

Theorem 3.2. *Let $n = m$ and $\alpha \ge 1/2$, $\beta, R, r, \zeta > 0$, $\gamma \ge 4$ and grant Assumption A with $\ell = \langle \alpha \rangle + 1$. Then the estimator $\widetilde{q}_\tau$ as defined in (13) with $\mathcal{B}_n$ from (9) satisfies uniformly over $f \in \mathcal{C}^\alpha(R, r, \zeta, U_n)$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$ as $n \to \infty$,*

$$|\widetilde{q}_\tau - q_\tau| = \mathcal{O}_P\big(\psi_{n(\delta \log n)^{-6}}(\alpha, \beta)\big),$$

where $\psi_\cdot(\alpha, \beta)$ is given in (5).

As the theorem shows, the adaptive method achieves the minimax rate up to a logarithmic factor. This additional loss is dominated by the stochastic error which is due to the estimation of φ_ε. Since Y_j and ε_k^* are not independent, we have to bound the stochastic error of $\widetilde{q}_{\tau,b}$ in a way that separates the error terms coming from the estimation of φ and φ_ε, respectively. Estimating the remaining parts, we lose the factor $(\delta \log n)^6$, which appears not to be optimal. To improve the rate slightly, $\delta = \delta_n$ could be chosen as a null sequence provided $\delta_n (\log n)^{1/2} \to \infty$. In the case where the error density is known, we can achieve the better rate $\psi_{n/\log\log n}(\alpha, \beta)$. The $\log\log n$-factor is the additional payment for $\mathcal{O}_P$-adaptivity, which is known to be unavoidable for a bounded loss function in standard regression, cf. Spokoiny [27]. For estimating the distribution function, an analogous result can be obtained, but is omitted.

4. Numerical results

4.1. Simulation study

We illustrate the implementation of the adaptive estimation procedure of Section 3. Our small simulation study serves as a proof of viability of the proposed method.

We run 1000 Monte Carlo simulations for four experimental setups. The sample size is set to $n = 1000$ and the external sample of the directly observed error is set to $m = 1000$ as well (here

Table 1. Empirical root mean square error (RMSE) of the adaptive deconvolution estimator and the empirical quantiles of (Y_j) (in parenthesis) for estimating q_τ based on 1000 Monte Carlo simulations with $n = m = 1000$

RMSE	$k = 1, \beta = 1$	$k = 2, \beta = 1$	$k = 1, \beta = 2$	$k = 2, \beta = 2$
$\tau = 0.1$	0.532 (0.886)	0.252 (0.706)	0.378 (1.029)	0.191 (0.765)
$\tau = 0.2$	0.265 (0.653)	0.114 (0.508)	0.175 (0.452)	0.091 (0.349)
$\tau = 0.3$	0.111 (0.461)	0.070 (0.360)	0.077 (0.178)	0.090 (0.158)
$\tau = 0.4$	0.067 (0.282)	0.080 (0.212)	0.112 (0.052)	0.105 (0.064)
$\tau = 0.5$	0.123 (0.110)	0.092 (0.096)	0.171 (0.175)	0.116 (0.145)
$\tau = 0.6$	0.162 (0.122)	0.094 (0.123)	0.200 (0.318)	0.109 (0.255)
$\tau = 0.7$	0.154 (0.326)	0.098 (0.272)	0.189 (0.462)	0.098 (0.373)
$\tau = 0.8$	0.107 (0.597)	0.150 (0.481)	0.115 (0.624)	0.141 (0.506)
$\tau = 0.9$	0.232 (1.015)	0.312 (0.783)	0.226 (0.849)	0.293 (0.675)

the external sample is independent of the main one). We consider $\Gamma(1, 1)$ and $\Gamma(2, 1)$ for the distribution of X where $\Gamma(k, \eta)$ denotes the gamma distribution with shape parameter k and scale η. Note that the shape k of the gamma distribution determines the Sobolev smoothness of the density while the density is smooth away from the origin. For the error distribution, we consider $\Gamma(1, \sqrt{2})$ centered around zero which corresponds to $\beta = 1$ and the standard Laplace distribution (scale equals 1) corresponding to $\beta = 2$. In both cases, the variance of the error equals 2.

The target quantiles of interest are q_τ with $\tau = 0.1, 0.2, \dots, 0.9$. In the real data example in the next subsection we compare the adaptive estimator to the "naive" quantile estimator given by the sample quantiles of the observations Y. Therefore we have also applied the naive estimator in the simulations. The results of this simulation study are given in Table 1. We can see that the results support the theory – the empirical root mean squared error (RMSE) is higher in most cases for $\beta = 2$ than for $\beta = 1$. Also, we can see that in most cases the RMSE is lower for $k = 2$ than for $k = 1$ since the gamma distribution with larger shape parameter is smoother in our context. At the tails, our estimation method is significantly better than the naive estimator. Near the median the naive estimator behaves nice when the distribution of the error is Laplace. This is not the case under the gamma error distribution which may suggest that the naive estimator profits from the symmetry of the error distribution. Similar behavior was observed also in distribution deconvolution with nonsymmetric error distributions, see Dattner and Reiser [5].

4.2. Real data example

High blood pressure is a direct cause of serious cardiovascular disease (Kannel *et al.* [15]) and determining reference values for physicians is important. In particular, estimating percentiles of systolic and diastolic blood pressure by sex, race or ethnicity, age, etc. is of substantial interest. Blood pressure is known to be measured with additional error which needs to be addressed in its analysis (see e.g., Frese, Fick and Sadowsky [9]). Therefore, measurement errors should be taken into account, otherwise quantile estimates based on the observed blood pressure measurements would be biased.

We illustrate our method using data from the Framingham Heart Study (Carroll *et al.* [2]). This study consists of a series of exams taken two years apart where systolic blood pressure (SBP) measurements of 1615 men aged 31–65 were taken. These data were used as an illustration for density deconvolution by Stirnemann, Comte and Samson [28] and for distribution deconvolution by Dattner and Reiser [5]. We denote by $Y_{j,1}$ and $Y_{j,2}$ the two repeated measures of SBP for each individual j at two different exams and denote by X_j the long-term average SBP of individual j. Then we model that

$$Y_{j,1} = X_j + \varepsilon_{j,1}, \qquad Y_{j,2} = X_j + \varepsilon_{j,2},$$

for individuals $j = 1, \dots, n$. Following Carroll *et al.* [2], we use the average of the two exams $Y_j' = (Y_{j,1} + Y_{j,2})/2$, so that the model in our case is

$$Y_j' = X_j + \varepsilon_j',$$

where $\varepsilon_j' = (\varepsilon_{j,1} + \varepsilon_{j,2})/2$.

Taking advantage of the repeated measurements, we can avoid parametric assumptions regarding the distribution of the errors. The only assumption we will make is that the distribution of the measurement error is symmetric around zero and does not vanish. We then set $\varepsilon_j^* = (Y_{j,1} - Y_{j,2})/2$ and note that under the symmetry assumption it is distributed as ε_j'. We emphasize the fact that our theoretical results do not require that the sample ε_j^* must be independent from that of the Y_j'.

Histograms of Y' and ε^* are presented in Figure 1. Although Figure 1 may suggest that the error distribution does not entirely satisfy the symmetry assumption, it serves as working hypothesis for our procedure and, indeed, it is supposed in previous works on the same data set as well. The resulting adaptive and naive quantiles estimates are displayed in Figure 2. We can see certain differences between the naive and adaptive estimates which might result in important implications for medical research, but here we do not aim at pursuing a more detailed statistical analysis.

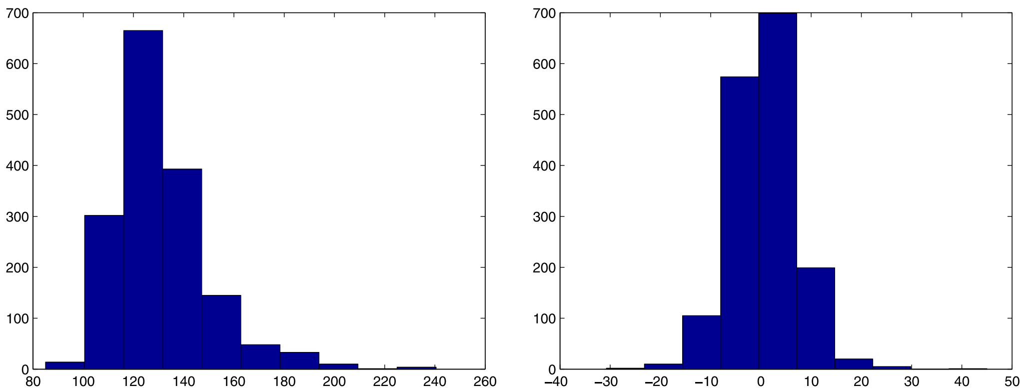

Figure 1. Average systolic blood pressure Y' (*left*) and the errors ε^* (*right*) over the two measurements from the two visits of 1615 men aged 31–65 from the Framingham Heart Study.

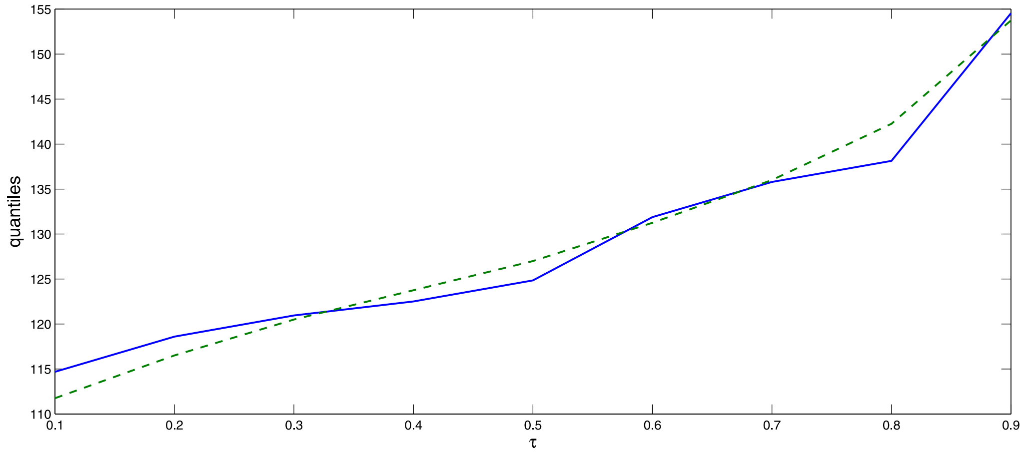

Figure 2. Quantiles estimates for systolic blood pressure of 1615 men aged 31–65 from the Framingham Heart Study. Solid line for the adaptive deconvolution estimator and dashed line for the empirical quantiles of (Y_j).

5. Proofs

5.1. Proofs for Section 2

For convenience, we will write $A_n(\vartheta) \lesssim B_n(\vartheta)$ if $A_n(\vartheta) = \mathcal{O}(B_n(\vartheta))$. For a better readability, we assume throughout $\beta \neq 1/2$. In the special case, $\beta = 1/2$ the order of the stochastic error will be $(\log n/n)^{1/2}$ which can be easily seen below in the bounds (24) and (26). The subscript n at the bandwidth will be omitted.

Since $1/\varphi_{\varepsilon,m}$ might explode for large stochastic errors we need the following lemma.

Lemma 5.1. *Suppose* $\mathbb{E}[|\varepsilon_k^*|^\delta] < \infty$ *for some* $\delta > 0$. *Let* $T_m \to \infty$ *be an increasing sequence satisfying* $m^{1/2} \inf_{u\in[-T_m,T_m]} |\varphi_\varepsilon(u)| \gtrsim (\log T_m)^2$, *then for any* $p < 2$

$$P\Big(\inf_{u\in[-T_m,T_m]} \big|\varphi_{\varepsilon,m}(u)\big| < m^{-1/2}(\log T_m)^p\Big) = \mathrm{o}(1) \qquad \text{as } m \to \infty.$$

Proof. The triangle inequality, the assumption on T_m and Markov's inequality yield for m as well as T_m large enough

$$\begin{aligned}
&P\Big(\inf_{u\in[-T_m,T_m]} \big|\varphi_{\varepsilon,m}(u)\big| < m^{-1/2}(\log T_m)^p\Big) \\
&\quad \le P\Big(\sup_{u\in[-T_m,T_m]} \big|\varphi_\varepsilon(u) - \varphi_{\varepsilon,m}(u)\big| > \inf_{u\in[-T_m,T_m]} \big|\varphi_\varepsilon(u)\big| - m^{-1/2}(\log T_m)^p\Big) \\
&\quad \lesssim \frac{2}{(\log T_m)^2} \mathbb{E}\Big[\sup_{u\in[-T_m,T_m]} m^{1/2} \big|\varphi_\varepsilon(u) - \varphi_{\varepsilon,m}(u)\big|\Big].
\end{aligned}$$

Noting $\mathbb{1}_{[-T_m,T_m]}(u) \le w(u)/w(T_m)$ for $w(u) := (\log(\mathrm{e}+|u|))^{-1/2-\eta}$ for some $\eta \in (0, 1/2)$, the above display can be bounded by

$$\frac{2}{w(T_m)(\log T_m)^2}\mathbb{E}\Big[\sup_{u\in\mathbb{R}} m^{1/2} w(u)\big|\varphi_\varepsilon(u) - \varphi_{\varepsilon,m}(u)\big|\Big] \lesssim (\log T_m)^{-3/2+\eta}, \qquad (14)$$

where the expectation is bounded by applying Theorem 4.1 in Neumann and Reiß [23]. □

To ensure consistency of the density estimator, the bandwidth satisfies usually $(n\wedge m)b^{2\beta+1} \to \infty$ and is of polynomial order in n, m. This implies $m^{1/2}\inf_{u\in[-1/b,1/b]}|\varphi_\varepsilon(u)| \gtrsim |\log b|^2$ for $f \in \mathcal{D}^\beta(R,\gamma)$, $\gamma > 0$, and thus Lemma 5.1 can be applied to $T_m = 1/b$. Under this conditions on b the probability of the event $B_\varepsilon(b)$, defined in (3), tends to one. In that case, it suffices to control terms on $B_\varepsilon(b)$, a strategy that will follow in the sequel. For instance, the $\mathcal{O}_P$-convergence in Theorem 2.7 is equivalent to $\lim_{C\to\infty}\lim_{n,m\to\infty} P(|\widetilde{q}_{\tau,b^*} - q_\tau| > C\psi_{n\wedge m}(\alpha,\beta)) = 0$ for which we have

$$\begin{aligned}
&\lim_{C\to\infty}\lim_{n,m\to\infty} P\big(|\widetilde{q}_{\tau,b^*} - q_\tau| > C\psi_{n\wedge m}(\alpha,\beta)\big)\\
&\quad\le \lim_{C\to\infty}\lim_{n,m\to\infty} P\big(\big\{|\widetilde{q}_{\tau,b^*} - q_\tau| > C\psi_{n\wedge m}(\alpha,\beta)\big\}\cap B_\varepsilon\big(b^*\big)\big) + \lim_{m\to\infty} P\big(B_\varepsilon\big(b^*\big)^c\big),
\end{aligned}$$

where the second term converges to zero by Lemma 5.1 and it remains to bound the first one.

On $B_\varepsilon(b)$ the weaker estimate $|\varphi_{\varepsilon,m}(u)| \ge m^{-1/2}$ for $|u| \le 1/b$ will frequently be enough, implying

$$\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\mathbb{1}_{\{|\varphi_{\varepsilon,m}(u)|\ge m^{-1/2}\}} = \frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)} \qquad \text{on } B_\varepsilon(b).$$

5.1.1. *Proof of Lemma* 2.1

On $B_\varepsilon(b)$, we have by continuity of the characteristic functions and the properties of the kernel that $g(u) := \frac{\varphi_n(u)\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}$ satisfies $g, g' \in L^2(\mathbb{R})$. Hence, $(1+x^2)^{1/2}\mathcal{F}^{-1}g(x) \in L^2(R)$ and the Cauchy–Schwarz inequality yields $\|\widetilde{f}_b\|_{L^1} \le \|(1+x^2)^{-1/2}\|_{L^2}\|(1+x^2)^{1/2}\mathcal{F}^{-1}g(x)\|_{L^2} < \infty$ on $B_\varepsilon(b)$. In particular, (2) is well defined on the event $B_\varepsilon(b)$.

On $B_\varepsilon(b)$, we have moreover $\lim_{\eta\to-\infty}\int_{-\infty}^{\eta}\widetilde{f}_b(x)\,\mathrm{d}x = 0$, by integrability of $\widetilde{f}_b$, and $\int_{-\infty}^{\infty}\widetilde{f}_b(x)\,\mathrm{d}x = \mathcal{F}[\widetilde{f}_b](0) = \varphi_n(0)\varphi_K(0)/\varphi_{\varepsilon,m}(0) = 1$. Applying $\|\widetilde{f}_b\|_\infty \le \|\varphi_K(bu)/\varphi_{\varepsilon_m}(u)\|_{L^1} < \infty$, we conclude that $\eta \mapsto \int_{-\infty}^{\eta}\widetilde{f}_b(x)\,\mathrm{d}x$ continuous and $[0,1]$ is contained in its range.

5.1.2. *Proof of Lemma* 2.2

Note that the assumption on φ_ε imply $|(\varphi_\varepsilon^{-1})'(u)| \lesssim (1+|u|)^{\beta-1}$ as well as $|\varphi_\varepsilon^{-1}(u)| \lesssim (1+|u|)^\beta$, $u \in \mathbb{R}$. We define the random Fourier multiplier

$$\psi(u) := (1+\mathrm{i}u)^{-\beta}\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}, \qquad u \in \mathbb{R}.$$

On $B_\varepsilon(b)$, as defined in (3), we will check Hörmander type conditions and derive an upper bound for the operator norm of $\psi(u)$. Hence, we have to determine a suitable constant $A_\psi > 0$ satisfying

$$\begin{aligned} \max_{l\in\{0,1\}} \Big(\int_{[-2,2]} \big|\psi^{(l)}(u)\big|^2 \,\mathrm{d}u\Big)^{1/2} &\le A_\psi \quad \text{and} \\ \max_{l\in\{0,1\}} \sup_{T\in[1,\infty)} T^{l-1/2}\Big(\int_{T\le|u|\le 4T} \big|\psi^{(l)}(u)\big|^2 \,\mathrm{d}u\Big)^{1/2} &\le A_\psi. \end{aligned} \tag{15}$$

To find A_ψ, we note that

$$\frac{1}{|\varphi_{\varepsilon,m}(u)|^p} \le \frac{p}{|\varphi_\varepsilon(u)|^p} + \frac{p|\varphi_{\varepsilon,m}(u)-\varphi_\varepsilon(u)|^p}{|\varphi_\varepsilon(u)\varphi_{\varepsilon,m}(u)|^p}, \qquad \text{for } p\in\{1,2\} \tag{16}$$

and thus on $B_\varepsilon(b)$

$$\frac{1}{|\varphi_{\varepsilon,m}(u)|} \le \frac{1+\Delta_m(u)}{|\varphi_\varepsilon(u)|}, \qquad \Delta_m(u) := \frac{m^{1/2}}{|\log b|^{3/2}}\big|\varphi_{\varepsilon,m}(u)-\varphi_\varepsilon(u)\big|.$$

By the assumptions on φ_ε and K, we conclude

$$\big|\psi(u)\big| \le \frac{|\varphi_K(bu)|(1+\Delta_m(u))}{(1+u^2)^{\beta/2}|\varphi_\varepsilon(u)|} \lesssim \big(1+\Delta_m(u)\big)\mathbb{1}_{[-1/b,1/b]}(u). \tag{17}$$

Concerning the derivative, we estimate $b \le 2(1+|u|)^{-1}$ for $|u|\le 1/b$ and $b<1/2$ and consequently by $|\varphi_\varepsilon'(u)/\varphi_\varepsilon(u)| \lesssim (1+|u|)^{-1}$

$$\begin{aligned} \big|\psi'(u)\big| &\le (\beta+1)\big(1+u^2\big)^{-(\beta+1)/2}\left|\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\right| + b\big(1+u^2\big)^{-\beta/2}\left|\frac{\varphi_K'(bu)}{\varphi_{\varepsilon,m}(u)}\right| \\ &\quad + \big(1+u^2\big)^{-\beta/2}\left|\frac{\varphi_{\varepsilon,m}'(u)}{\varphi_{\varepsilon,m}(u)}\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\right| \\ &\lesssim \frac{|\psi(u)|}{1+|u|} + \big|\psi(u)\big|\left|\frac{\varphi_{\varepsilon,m}'(u)}{\varphi_{\varepsilon,m}(u)}\right| \\ &\lesssim \big(1+\Delta_m(u)\big)\left(\frac{1}{1+|u|} + \big(1+\Delta_m(u)\big)\left|\frac{\varphi_{\varepsilon,m}'(u)}{\varphi_\varepsilon(u)}\right|\right)\mathbb{1}_{[-1/b,1/b]}(u) \\ &\lesssim \big(1+\Delta_m(u)\big)\left(\frac{2+\Delta_m(u)}{1+|u|} + \big(1+\Delta_m(u)\big)\big(1+|u|\big)^\beta\big|\varphi_{\varepsilon,m}'(u)-\varphi_\varepsilon'(u)\big|\right) \\ &\quad \times \mathbb{1}_{[-1/b,1/b]}(u) \\ &\lesssim \frac{(1+\Delta_m(u))^2}{1+|u|}\big(1+\big(1+|u|\big)^{\beta+1}\big|\varphi_{\varepsilon,m}'(u)-\varphi_\varepsilon'(u)\big|\big)\mathbb{1}_{[-1/b,1/b]}(u). \end{aligned} \tag{18}$$

With these bounds at hand, we can show (15). For $l=0$, the estimate (17) and $1/T \lesssim (1+|u|)^{-1}$ for $|u| \le 4T$ yield

$$\int_{-2}^{2} \left|\psi(u)\right|^2 \mathrm{d}u \lesssim \int_{-2}^{2} \left(1+\Delta_m^2(u)\right)\mathbb{1}_{[-1/b,1/b]}(u)\,\mathrm{d}u,$$

$$\begin{aligned}\frac{1}{T}\int_{T\le|u|\le 4T} \left|\psi(u)\right|^2 \mathrm{d}u &\lesssim \frac{1}{T}\int_{T\le|u|\le 4T} \left(1+\Delta_m^2(u)\right)\mathbb{1}_{[-1/b,1/b]}(u)\,\mathrm{d}u \\ &\lesssim 1+\int_{-1/b}^{1/b}\left(1+|u|\right)^{-1}\Delta_m^2(u)\,\mathrm{d}u,\end{aligned}$$

for b small enough. Hence, the conditions (15) for $l=0$ are satisfied for A_ψ of the order $(1+\int_{-1/b}^{1/b}(1+|u|)^{-1}\Delta_m^2(u)\,\mathrm{d}u)^{1/2}$. For $l=1$, we verify by (18) and $T \le (1+|u|)$ for $|u|>T$

$$\int_{-2}^{2} \left|\psi'(u)\right|^2 \mathrm{d}u \lesssim \int_{-2}^{2} \left(1+\Delta_m^4(u)\right)\left(1+\left(1+|u|\right)^{2\beta+2}\left|\varphi'_{\varepsilon,m}(u)-\varphi'_{\varepsilon}(u)\right|^2\right)\mathrm{d}u \quad \text{and}$$

$$\begin{aligned}&T\int_{T\le|u|\le 4T}\left|\psi'(u)\right|^2\mathrm{d}u \\ &\lesssim \int_{T\le|u|\le 4T}\frac{T\,\mathrm{d}u}{(1+|u|)^2} \\ &\quad+\int_{-1/b}^{1/b}\left(\frac{\Delta_m^4(u)}{1+|u|}+\left(1+\Delta_m^4(u)\right)\left(1+|u|\right)^{2\beta+1}\left|\varphi'_{\varepsilon,m}(u)-\varphi'_{\varepsilon}(u)\right|^2\right)\mathrm{d}u \\ &\lesssim 1+\int_{-1/b}^{1/b}\left(\frac{\Delta_m^4(u)}{1+|u|}+\left(1+\Delta_m^4(u)\right)\left(1+|u|\right)^{2\beta+1}\left|\varphi'_{\varepsilon,m}(u)-\varphi'_{\varepsilon}(u)\right|^2\right)\mathrm{d}u.\end{aligned}$$

Therefore, we find a constant $A'>0$, depending only on R,β, such that (15) holds for

$$\begin{aligned}A_\psi := A'\Bigg(1+\int_{-1/b}^{1/b}\bigg(&\frac{\Delta_m^2(u)+\Delta_m^4(u)}{1+|u|} \\ &+\left(1+\Delta_m^4(u)\right)\left(1+|u|\right)^{2\beta+1}\left|\varphi'_{\varepsilon,m}(u)-\varphi'_{\varepsilon}(u)\right|^2\bigg)\mathrm{d}u\Bigg)^{1/2}.\end{aligned} \tag{19}$$

The conditions (15) imply that ψ is indeed a Fourier multiplier on $B_\varepsilon(b)$ and thus by Theorem 4.8 and Corollary 4.13 by Girardi and Weis [10] with $p=2$, $l=1$ there is a universal constant $C>0$ such that for all $\eta>0$ and $f\in C^{s+\beta+\eta}(\mathbb{R})$

$$\left\|\mathcal{F}^{-1}\left[\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\right]*f\right\|_{C^s}=\left\|\mathcal{F}^{-1}\left[\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\mathcal{F}f\right]\right\|_{C^s}\le CA_\psi\left\|\mathcal{F}^{-1}\left[(1+\mathrm{i}u)^\beta\mathcal{F}f\right]\right\|_{C^{s+\eta}}.$$

Choosing $\eta>0$ such that $s+\beta+\eta, s+\eta\notin\mathbb{N}$, the Fourier multiplier $(1+\mathrm{i}u)^\beta$ induces an isomorphism from $C^{s+\beta+\eta}(\mathbb{R})$ onto $C^{s+\eta}(\mathbb{R})$ (Triebel [29], Thm. 2.3.8). Hence, there is another universal constant $C' > 0$ such that the second assertion of the lemma follows:

$$\left\| \mathcal{F}^{-1}\left[\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\mathcal{F}f\right]\right\|_{C^s} \le \mathcal{E}_b \|f\|_{C^{s+\beta+\eta}} \qquad \text{with } \mathcal{E}_b := C'A_\psi.$$

To bound $\mathcal{E}_b$, we apply Markov's inequality on A_ψ from (19). The inequality by Rosenthal [25] yields

$$\sup_{u\in\mathbb{R}} \mathbb{E}\big[m^{p/2}\big|\varphi_{\varepsilon,m}^{(l)}(u) - \varphi_\varepsilon^{(l)}(u)\big|^p\big] < \infty$$

for $l = 0$ and $p \in \mathbb{N}$ as well as $l = 1$ and $p \in \{1, \dots, 4\}$. Combined with the Markov inequality and Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}
&P\left(B_\varepsilon(b) \cap \left\{\mathcal{E}_b > \frac{c^{1/2}}{m^{1/2}b^{\beta+1} \wedge 1}\right\}\right) \\
&\quad \le c^{-1}\big(mb^{2\beta+2} \wedge 1\big)\mathbb{E}\big[\mathcal{E}_b^2 \mathbb{1}_{B_\varepsilon(b)}\big] \\
&\quad \lesssim \frac{1}{c}\big(mb^{2\beta+2} \wedge 1\big)\left(1 + \int_{-1/b}^{1/b} \Big(\big(1+|u|\big)^{-1}\mathbb{E}\big[\Delta_m^2(u) + \Delta_m^4(u)\big]\right. \qquad (20)\\
&\qquad\qquad \left. + \mathbb{E}\Big[\big(1+\Delta_m^4(u)\big)\big(1+|u|\big)^{2\beta+1}\big|\varphi_{\varepsilon,m}'(u) - \varphi_\varepsilon'(u)\big|^2\Big]\Big)\,\mathrm{d}u\right) \\
&\quad \lesssim \frac{mb^{2\beta+2} \wedge 1}{c}\left(1 + \frac{1}{|\log b|^3}\int_{-1/b}^{1/b} \frac{\mathrm{d}u}{1+|u|} + \frac{1}{m}\int_{-1/b}^{1/b}\big(1+|u|\big)^{2\beta+1}\,\mathrm{d}u\right) \lesssim \frac{1}{c},
\end{aligned}$$

which shows $\mathcal{E}_b = \mathcal{O}_P(m^{-1/2}b^{-\beta-1} \vee 1)$.

5.1.3. *Proof of Proposition* 2.5

The following lemma establishes a bound for the bias term of the estimator for the distribution function.

Lemma 5.2. *Let Assumption* A *hold with* $\ell = \langle\alpha\rangle + 1$, $\alpha > 0$ *and* $f(\cdot + q_\tau) \in C^\alpha([-\zeta, \zeta], R)$. *Then we have*

$$\sup_{f(\cdot+q_\tau)\in C^\alpha([-\zeta,\zeta],R)} \left|\int_{-\infty}^{q_\tau} K_b * f(x)\,\mathrm{d}x - \int_{-\infty}^{q_\tau} f(x)\,\mathrm{d}x\right| \le Db^{\alpha+1},$$

where $D = (R/(\langle\alpha\rangle + 1)! + 2\zeta^{-\alpha-1})\|K(x)x^{\alpha+1}\|_{L^1}$.

Proof. Let $F(x) := \int_{-\infty}^x f(y)\,\mathrm{d}y$. Fubini's theorem yields

$$\int_{-\infty}^{q_\tau} K_b * f(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} K_b(x)F(q_\tau - x)\,\mathrm{d}x,$$

where $K_b(x) := b^{-1}K(x/b)$, $x \in \mathbb{R}$. Therefore, the bias depends only locally on f. Note that $F(\cdot + q_\tau) \in C^{\alpha+1}([-\zeta, \zeta])$ by assumption. A Taylor expansion of F around q_τ yields for $|bz| < \zeta$

$$F(q_\tau - bz) - F(q_\tau) = -bzF'(q_\tau) + \cdots + (-bz)^{\langle\alpha\rangle+1}\frac{F^{(\langle\alpha\rangle+1)}(q_\tau - \kappa bz)}{(\langle\alpha\rangle + 1)!},$$

where $0 \le \kappa \le 1$. Using the fact that $\int x^k K(x)\,\mathrm{d}x = 0$ for $k = 1, \ldots, \langle\alpha\rangle + 1$ and the properties of the class, we obtain

$$\begin{aligned}
&\left|\int_{-\infty}^{q_\tau}\bigl(K_b * f(x) - f(x)\bigr)\,\mathrm{d}x\right| \\
&\quad= \left|\int_{-\infty}^{\infty} K(z)\bigl(F(q_\tau - bz) - F(q_\tau)\bigr)\,\mathrm{d}z\right| \\
&\quad\le \left|\int_{|z|<\zeta/b} K(z)(-bz)^{\langle\alpha\rangle+1}\frac{F^{(\langle\alpha\rangle+1)}(q_\tau - \kappa bz) - F^{(\langle\alpha\rangle+1)}(q_\tau)}{(\langle\alpha\rangle + 1)!}\,\mathrm{d}z\right| \\
&\qquad+ \int_{|z|\ge\zeta/b}\bigl|K(z)\bigr|\bigl|F(q_\tau - bz) - F(q_\tau)\bigr|\,\mathrm{d}z \\
&\quad\le \frac{b^{\langle\alpha\rangle+1}R}{(\langle\alpha\rangle + 1)!}\int_{-\infty}^{\infty}\bigl|K(z)\bigr||z|^{\langle\alpha\rangle+1}|\kappa bz|^{\alpha+1-(\langle\alpha\rangle+1)}\,\mathrm{d}z + 2\int_{|z|\ge\zeta/b}\bigl|K_b(z)\bigr|\,\mathrm{d}z \\
&\quad\le \left(\frac{b^{\alpha+1}R}{(\langle\alpha\rangle + 1)!} + 2\left(\frac{b}{\zeta}\right)^{\alpha+1}\right)\int_{-\infty}^{\infty}\bigl|K(z)\bigr||z|^{\alpha+1}\,\mathrm{d}z,
\end{aligned}$$

and the statement follows. □

Proof of Proposition 2.5. We will show uniformly over $f \in \mathcal{C}^\alpha(R, r, \zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$ for any b such that $(n \wedge m)b^{2\beta+1} \to \infty$

$$\begin{aligned}
&\left|\int_{-\infty}^{q_\tau}\bigl(\widetilde{f}_b(x) - f(x)\bigr)\,\mathrm{d}x\right| \\
&\quad= \mathcal{O}_P\left(b^{\alpha+1} + \frac{1}{\sqrt{(n \wedge m)(b^{2\beta-1} \wedge 1)}} + \frac{1}{\sqrt{(n \wedge m)(mb^{2\beta+2} \wedge 1)}}\right).
\end{aligned}$$

The third term on the right-hand side is of smaller or of the same order than the second one if and only if $(mb^{1\wedge 2\beta+2})^{-1} \lesssim 1$. Hence, when $\alpha \ge 1/2$ the asymptotically optimal choice $b = (n \wedge m)^{-1/(2\alpha+2(\beta\vee 1/2)+1)}$ yields

$$\left|\int_{-\infty}^{q_\tau}\bigl(\widetilde{f}_b(x) - f(x)\bigr)\,\mathrm{d}x\right| = \mathcal{O}_P\bigl((n \wedge m)^{-(\alpha+1)/(2\alpha+2\beta+1)} \vee (n \wedge m)^{-1/2}\bigr).$$

Step 1: As usual, we decompose the error into a deterministic error term and a stochastic error term, writing $\varphi_X = \mathcal{F} f$,

$$\begin{aligned}&\left|\int_{-\infty}^{q_\tau}\big(\widetilde{f}_b(x) - f(x)\big)\,\mathrm{d}x\right|\\&\quad\le \left|\int_{-\infty}^{q_\tau}\big(K_b * f(x) - f(x)\big)\,\mathrm{d}x\right| + \left|\int_{-\infty}^{q_\tau}\mathcal{F}^{-1}\left[\frac{\varphi_n(u)\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)} - \varphi_K(bu)\varphi_X(u)\right](x)\,\mathrm{d}x\right|.\end{aligned}$$

The bias is of order $\mathcal{O}(b^{\alpha+1})$ by Lemma 5.2. As discussed above, we decompose the stochastic error into a singular part and a continuous one using a smooth truncation function. Let $a_c \in C^\infty(\mathbb{R})$ satisfy $a_c(x) = 1$ for $x \le -1$ and $a_c(x) = 0$ for $x \ge 0$ and define $a_s(x) := \mathbb{1}_{(-\infty,0]}(x) - a_c(x)$. Then

$$\begin{aligned}&\int_{-\infty}^{q_\tau}\mathcal{F}^{-1}\left[\varphi_K(bu)\left(\frac{\varphi_n(u)}{\varphi_{\varepsilon,m}(u)} - \varphi_X(u)\right)\right](x)\,\mathrm{d}x\\&\quad= \int_{\mathbb{R}} a_s(x)\mathcal{F}^{-1}\left[\varphi_K(bu)\left(\frac{\varphi_n(u)}{\varphi_{\varepsilon,m}(u)} - \varphi_X(u)\right)\right](x+q_\tau)\,\mathrm{d}x\\&\qquad+ \int_{\mathbb{R}} a_c(x)\mathcal{F}^{-1}\left[\varphi_K(bu)\left(\frac{\varphi_n(u)}{\varphi_{\varepsilon,m}(u)} - \varphi_X(u)\right)\right](x+q_\tau)\,\mathrm{d}x\\&\quad=: T_s + T_c.\end{aligned} \tag{21}$$

The singular term T_s will be treated in the next step while we bound the continuous, but not integrable term T_c in Step 3.

Step 2: Lemma 5.1 shows that the probability of the complement $B_\varepsilon(b)^c$ of $B_\varepsilon(b)$ from (3) converges to zero. We obtain for any $c > 0$ with Markov's inequality

$$\begin{aligned}&P\left(|T_s| > \frac{c}{\sqrt{(n\wedge m)(b^{2\beta-1}\vee 1)}}\right)\\&\quad\le P\left(B_\varepsilon(b)\cap\left\{|T_s| > \frac{c}{\sqrt{(n\wedge m)(b^{2\beta-1}\vee 1)}}\right\}\right) + P\big(B_\varepsilon(b)^c\big)\\&\quad\le \frac{1}{c}\sqrt{(n\wedge m)\big(b^{2\beta-1}\vee 1\big)}\,\mathbb{E}\big[|T_s|\mathbb{1}_{B_\varepsilon(b)}\big] + \mathrm{o}(1).\end{aligned}$$

To bound $\mathbb{E}[|T_s|\mathbb{1}_{B_\varepsilon(b)}]$, we first note by Plancherel's identity

$$\begin{aligned}T_s &= \frac{1}{2\pi}\int_{\mathbb{R}}\mathcal{F}a_s(u)\mathrm{e}^{-\mathrm{i}uq_\tau}\varphi_K(bu)\left(\frac{\varphi_n(u)}{\varphi_{\varepsilon,m}(u)} - \varphi_X(u)\right)\mathrm{d}u\\&= \frac{1}{2\pi}\int_{\mathbb{R}}\mathcal{F}a_s(u)\mathrm{e}^{-\mathrm{i}uq_\tau}\varphi_K(bu)\left(\frac{\varphi_n(u)}{\varphi_{\varepsilon}(u)} - \varphi_X(u)\right)\mathrm{d}u\\&\quad+ \frac{1}{2\pi}\int_{\mathbb{R}}\mathcal{F}a_s(u)\mathrm{e}^{-\mathrm{i}uq_\tau}\frac{\varphi_K(bu)\varphi_n(u)}{\varphi_\varepsilon(u)}\left(\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1\right)\mathrm{d}u\\&=: \frac{1}{2\pi}(T_{s,x} + T_{s,\varepsilon}).\end{aligned} \tag{22}$$

The first term, $T_{s,x}$ corresponds to the error due to the unknown density f while $T_{s,\varepsilon}$ is dominated by the error of the estimator $\varphi_{\varepsilon,m}$. Since a_s is of bounded variation and has compact support, there is a constant $A_s \in (0,\infty)$ such that $|\mathcal{F}a_s(u)| \le A_s(1+|u|)^{-1}$. Plancherel's identity yields

$$\begin{aligned}
\operatorname{Var}(T_{s,x}) = \mathbb{E}\big[|T_{s,x}|^2\big] &\le \frac{1}{n}\mathbb{E}\bigg[\bigg|\int_{\mathbb{R}} \mathcal{F}a_s(u)\mathrm{e}^{-\mathrm{i}uq_\tau}\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)}\mathrm{e}^{\mathrm{i}uY_1}\,\mathrm{d}u\bigg|^2\bigg] \\
&\le \frac{4\pi^2}{n}\|f_Y\|_\infty \bigg\|\mathcal{F}^{-1}\bigg[\mathcal{F}a_s(u)\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)}\bigg]\bigg\|_{L^2}^2 \\
&\le \frac{4\pi^2}{n}\|K\|_{L^1}^2\|f_Y\|_\infty\int_{-1/b}^{1/b}\frac{|\mathcal{F}a_s(u)|^2}{|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u \\
&\le \frac{4\pi^2}{n}\|K\|_{L^1}^2 A_s^2\|f_Y\|_\infty\int_{-1/b}^{1/b}\frac{1}{(1+|u|)^2|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u.
\end{aligned} \tag{23}$$

Using the assumption $\|f\|_\infty < R$ and $f_\varepsilon \in \mathcal{D}^\beta(R,\gamma)$, we get

$$\mathbb{E}\big[|T_{s,x}|^2\big] \lesssim \frac{1}{n}\int_{-1/b}^{1/b}\big(1+|u|\big)^{2\beta-2}\,\mathrm{d}u \lesssim \frac{1}{nb^{2\beta-1}} \vee \frac{1}{n}. \tag{24}$$

To bound $T_{s,\varepsilon}$, we will use the following version of a lemma by Neumann [21]: by the definition (3) of $B_\varepsilon(b)$ and applying (16) it holds

$$\begin{aligned}
&\mathbb{E}\bigg[\bigg|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)}-1\bigg|^2\mathbb{1}_{B_\varepsilon(b)}\bigg] \\
&\quad\le 2\mathbb{E}\bigg[\frac{|\varphi_{\varepsilon,m}(u)-\varphi_\varepsilon(u)|^2}{|\varphi_\varepsilon(u)|^2}\bigg] + 2\mathbb{E}\bigg[\frac{|\varphi_{\varepsilon,m}(u)-\varphi_\varepsilon(u)|^4}{|\varphi_\varepsilon(u)\varphi_{\varepsilon,m}(u)|^2}\mathbb{1}_{B_\varepsilon(b)}\bigg] \\
&\quad\le \frac{2\mathbb{E}[|\varphi_{\varepsilon,m}(u)-\varphi_\varepsilon(u)|^2]}{|\varphi_\varepsilon(u)|^2} + \frac{2m\mathbb{E}[|\varphi_{\varepsilon,m}(u)-\varphi_\varepsilon(u)|^4]}{|\varphi_\varepsilon(u)|^2} \\
&\quad\le \frac{18}{m|\varphi_\varepsilon(u)|^2}.
\end{aligned} \tag{25}$$

We estimate with the Cauchy–Schwarz inequality

$$\begin{aligned}
T_{s,\varepsilon}^2 &\le \|K\|_{L^1}^2\int_{-1/b}^{1/b}\bigg|\frac{\varphi_n(u)}{\varphi_\varepsilon(u)}\bigg|^2\,\mathrm{d}u\int_{-1/b}^{1/b}\big|\mathcal{F}a_s(u)\big|^2\bigg|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)}-1\bigg|^2\,\mathrm{d}u \\
&\le 2\|K\|_{L^1}^2\bigg(\|\varphi_X\|_{L^2}^2 + \int_{-1/b}^{1/b}\frac{|\varphi_n(u)-\varphi_Y(u)|^2}{|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u\bigg) \\
&\quad\times\int_{-1/b}^{1/b}\big|\mathcal{F}a_s(u)\big|^2\bigg|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)}-1\bigg|^2\,\mathrm{d}u.
\end{aligned}$$

Applying again the Cauchy–Schwarz inequality, Fubini's theorem, the decay of $\mathcal{F}a_s$ and (25), we obtain

$$\begin{aligned}
&\mathbb{E}\big[|T_{s,\varepsilon}|\mathbb{1}_{B_\varepsilon(b)}\big] \\
&\quad\le \sqrt{2}\|K\|_{L^1}\bigg(\|\varphi_X\|_{L^2}^2 + \int_{-1/b}^{1/b}\frac{\mathbb{E}[|\varphi_n(u)-\varphi_Y(u)|^2]}{|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u\bigg)^{1/2} \\
&\qquad\times\bigg(\int_{-1/b}^{1/b}\frac{A_s^2}{(1+|u|)^2}\mathbb{E}\bigg[\bigg|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)}-1\bigg|^2\mathbb{1}_{B_\varepsilon(b)}\bigg]\,\mathrm{d}u\bigg)^{1/2} \\
&\quad\le \frac{\sqrt{36}\|K\|_{L^1}A_s}{\sqrt{m}}\bigg(\|\varphi_X\|_{L^2}^2+\int_{-1/b}^{1/b}\frac{\mathrm{d}u}{n|\varphi_\varepsilon(u)|^2}\bigg)^{1/2}\bigg(\int_{-1/b}^{1/b}\frac{\mathrm{d}u}{(1+|u|)^2|\varphi_\varepsilon(u)|^2}\bigg)^{1/2}.
\end{aligned} \tag{26}$$

The assumptions $\|f\|_\infty \lesssim 1$, $|\varphi_\varepsilon(u)| \lesssim (1+|u|)^{-\beta}$ and $n^{-1}b^{-2\beta-1}\to 0$ for the optimal $b=b^*$ yield

$$\mathbb{E}\big[|T_{s,\varepsilon}|\mathbb{1}_{B_\varepsilon(b)}\big] \lesssim \bigg(1+\frac{1}{nb^{2\beta+1}}\bigg)^{1/2}\bigg(\frac{1}{\sqrt{m}b^{\beta-1/2}}\vee\frac{1}{\sqrt{m}}\bigg) \lesssim \frac{1}{\sqrt{m}b^{\beta-1/2}}\vee\frac{1}{\sqrt{m}}.$$

Together with (24) and (22) this implies the optimal order

$$\mathbb{E}\big[|T_s|\mathbb{1}_{B_\varepsilon(b)}\big] \lesssim \big((n\wedge m)\big(b^{2\beta-1}\wedge 1\big)\big)^{-1/2}.$$

Step 3: The empirical measures of (Y_j) and (ε_k) are given by $\mu_{Y,n} := \frac{1}{n}\sum_{j=1}^n \delta_{Y_j}$ and $\mu_{\varepsilon,m} := \frac{1}{m}\sum_{k=1}^m \delta_{\varepsilon_k}$, respectively, with Dirac measure δ_x in $x\in\mathbb{R}$. We can write

$$\begin{aligned}
T_c &= \int_{\mathbb{R}} a_c(x)\mathcal{F}^{-1}\bigg[\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\big(\varphi_n(u)-\varphi_{\varepsilon,m}(u)\varphi_X(u)\big)\bigg](x+q_\tau)\,\mathrm{d}x \\
&= \mathcal{F}^{-1}\bigg[\frac{\varphi_K(-bu)}{\varphi_{\varepsilon,m}(-u)}\big(\varphi_n(-u)-\varphi_{\varepsilon,m}(-u)\varphi_X(-u)\big)\bigg]*a_c(-q_\tau) \\
&= \mathcal{F}^{-1}\bigg[\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\bigg]*\big(\mu_{Y,n}*a_c(-\bullet)-\mu_{\varepsilon,m}*f*a_c(-\bullet)\big)(q_\tau).
\end{aligned}$$

Applying Lemma 2.2, we obtain on $B_\varepsilon(b)$ for any integer $s>\beta$

$$\begin{aligned}
|T_c| &\le \bigg\|\mathcal{F}^{-1}\bigg[\frac{\varphi_K(bu)}{\varphi_{\varepsilon,m}(u)}\bigg]*\big(\mu_{Y,n}*a_c(-\bullet)-\mu_{\varepsilon,m}*f*a_c(-\bullet)\big)\bigg\|_\infty \\
&\le \mathcal{E}_b\big\|\mu_{Y,n}*a_c(-\bullet)-\mu_{\varepsilon,m}*f*a_c(-\bullet)\big\|_{C^s} \\
&\lesssim \mathcal{E}_b\sum_{l=0}^{s}\big\|\mu_{Y,n}*a_c^{(l)}(-\bullet)-\mu_{\varepsilon,m}*f*a_c^{(l)}(-\bullet)\big\|_\infty.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&P\biggl(B_\varepsilon(b)\cap\biggl\{|T_c|>\frac{c}{\sqrt{(n\wedge m)}(\sqrt{m}b^{\beta+1}\wedge 1)}\biggr\}\biggr)\\
&\quad\le P\biggl(B_\varepsilon(b)\cap\biggl\{\mathcal{E}_b>\biggl(\frac{c}{mb^{2\beta+2}\wedge 1}\biggr)^{1/2}\biggr\}\biggr)\\
&\qquad+P\Biggl(\sum_{l=0}^{s}\bigl\|\mu_{Y,n}*a_c^{(l)}-\mu_{\varepsilon,m}*f*a_c^{(l)}\bigr\|_\infty>\biggl(\frac{c}{n\wedge m}\biggr)^{1/2}\Biggr)\\
&\quad=:P_1+P_2.
\end{aligned}$$

By Lemma 2.2, more precisely estimate (20), the first probability is of the order $1/c$. To bound P_2, it suffices to show $\|\mu_{Y,n}*a_c^{(l)}-\mu_{\varepsilon,m}*f*a_c^{(l)}\|_\infty=\mathcal{O}_P((n\wedge m)^{-1/2})$ for all $l=0,\dots,s$. Denoting the density of Y_j as $f_Y=f*f_\varepsilon$, we decompose

$$\begin{aligned}
&\bigl\|\mu_{Y,n}*\bigl(a_c^{(l)}(-\bullet)\bigr)-\mu_{\varepsilon,m}*f*\bigl(a_c^{(l)}(-\bullet)\bigr)\bigr\|_\infty\\
&\quad\le\bigl\|\mu_{Y,n}*\bigl(a_c^{(l)}(-\bullet)\bigr)-f_Y*\bigl(a_c^{(l)}(-\bullet)\bigr)\bigr\|_\infty\\
&\qquad+\bigl\|f_\varepsilon*\bigl(f*\bigl(a_c^{(l)}(-\bullet)\bigr)\bigr)-\mu_{\varepsilon,m}*\bigl(f*\bigl(a_c^{(l)}(-\bullet)\bigr)\bigr)\bigr\|_\infty\\
&\quad\le\biggl\|\int a_c^{(l)}(y-\bullet)\mu_{Y,n}(\mathrm{d}y)-\mathbb{E}\bigl[a_c^{(l)}(Y_1-\bullet)\bigr]\biggr\|_\infty\\
&\qquad+\biggl\|\mathbb{E}\bigl[\bigl(f*a_c^{(l)}\bigr)(\varepsilon_1-\bullet)\bigr]-\int\bigl(f*a_c^{(l)}\bigr)(z-\bullet)\mu_{\varepsilon,m}(\mathrm{d}z)\biggr\|_\infty.
\end{aligned}$$

By construction all $a_c^{(l)}, l\ge 1$, have compact support and are bounded. Therefore, $\|a_c^{(l)}\|_{L^1}<\infty$, $\|(a_c*f)^{(l)}\|_{L^1}\le\|a_c^{(l)}\|_{L^1}\|f\|_{L^1}<\infty$ and thus $a_c^{(l)}(\bullet-t)$ and $a_c^{(l)}*f(\bullet-t)$, $l\ge 0$, are of bounded variation for all $t\in\mathbb{R}$. Since the set of functions with bounded variation is a Donsker class (cf. Theorem 2.1 by Dudley [7]), the two terms in the previous display converge in probability to a tight limit with $\sqrt{n}$-rate and $\sqrt{m}$-rate, respectively. Consequently,

$$\sqrt{n\wedge m}\bigl\|\mu_{Y,n}*\bigl(a_c^{(l)}(-\bullet)\bigr)-\mu_{\varepsilon,m}*f*\bigl(a_c^{(l)}(-\bullet)\bigr)\bigr\|_\infty=\mathcal{O}_P(1)$$

for all $\ell=0,\dots,s$ and P_2 is arbitrary small for c large. □

For the adaptive estimator, we will later need the following uniform version of Proposition 2.5.

Corollary 5.3. *Suppose Assumption A holds with $l=\langle\alpha\rangle+1$ and let the set $\mathcal{B}=\mathcal{B}_n$ be given by (9). For critical values $(\delta_b)_{b\in\mathcal{B}}$ satisfying $\delta_b>3Db^{\alpha+1}$ and for any sequence $(x_n)_n$ with $x_n\to\infty$ arbitrarily slowly we obtain uniformly in $\mathcal{C}^\alpha(R,r,\zeta)$ and $\mathcal{D}^\beta(R,\gamma)$*

$$\begin{aligned}
&P\biggl(\exists b\in\mathcal{B}\colon\biggl|\int_{-\infty}^{q_\tau}\bigl(\widetilde{f}_b(x)-f(x)\bigr)\,\mathrm{d}x\biggr|>\delta_b\biggr)\\
&\quad=\mathcal{O}\biggl(\sum_{b\in\mathcal{B}}\biggl(\frac{1}{\delta_b}\bigl((n\wedge m)\bigl(b^{2\beta-1}\wedge 1\bigr)\bigr)^{-1/2}+\frac{1}{\delta_b^2}\frac{x_n}{(n\wedge m)(mb^{2\beta+2}\wedge 1)}\biggr)\biggr)+\mathrm{o}(1).
\end{aligned}$$

In particular, if $|\mathcal{B}| \lesssim \log n$, $\max_{b\in\mathcal{B}} b \to 0$ *and* $\min_{b\in\mathcal{B}}(n \wedge m)b^{2\beta+1} \to \infty$, *then*

$$\sup_{b\in\mathcal{B}}\left|\int_{-\infty}^{q_\tau}\big(\widetilde{f}_b(x) - f(x)\big)\,\mathrm{d}x\right| \xrightarrow{P} 0.$$

Proof. With the notation of the proof of Proposition 2.5 and applying Lemma 5.2, we obtain

$$\begin{aligned}\left|\int_{-\infty}^{q_\tau}\big(\widetilde{f}_b(x) - f(x)\big)\,\mathrm{d}x\right| &\le \left|\int_{-\infty}^{q_\tau}\big(K_b * f(x) - f(x)\big)\,\mathrm{d}x\right| + |T_s| + |T_c| \\ &\le Db^{\alpha+1} + |T_s| + |T_c|,\end{aligned}$$

where T_s and T_c are the stochastic errors of the singular part and of the continuous part, respectively, as defined in (21). Since both terms depend on b let us write $T_s(b)$ and $T_c(b)$. By definition $b_1 \le b$ implies $B_\varepsilon(b_1) \subseteq B_\varepsilon(b)$. Then, Step 2 in the previous proof shows

$$\begin{aligned}P(\exists b \in \mathcal{B}\colon\ T_s > \delta_b/3) &\le \left(\sum_{b\in\mathcal{B}} P\big(\{T_s(b) > \delta_b/3\} \cap B_\varepsilon(b_1)\big)\right) + \mathrm{o}(1) \\ &\le \left(\sum_{b\in\mathcal{B}} \delta_b^{-1}\mathbb{E}\big[|T_s(b)|\mathbb{1}_{B_\varepsilon(b_1)}\big]\right) + \mathrm{o}(1) \\ &\lesssim \left(\sum_{b\in\mathcal{B}} \delta_b^{-1}\big((n\wedge m)\big(b^{2\beta-1}\wedge 1\big)\big)^{-1/2}\right) + \mathrm{o}(1).\end{aligned}$$

Following Step 3 in the previous proof, we obtain with the random operator norm $\mathcal{E}_b$, for some integer $s > \beta$ and for a diverging sequence $(x_{(n\wedge m)})$

$$\begin{aligned}&P(\exists b \in \mathcal{B}\colon\ T_c > \delta_b/3) \\ &\quad\le P\big(\big\{\exists b \in \mathcal{B}\colon\ \mathcal{E}_b > \delta_b(n\wedge m)^{1/2}/\big(3(x_{(n\wedge m)})^{1/2}\big)\big\} \cap B_\varepsilon(b_1)\big) + P\big(B_\varepsilon(b_1)^c\big) \\ &\qquad + P\left(\left\{\sum_{l=0}^{s}\big\|\mu_{Y,n} * a_c^{(l)} - \mu_{\varepsilon,m} * f * a_c^{(l)}\big\|_\infty > \left(\frac{x_{(n\wedge m)}}{(n\wedge m)}\right)^{1/2}\right\}\right) \\ &\quad\le \left(\sum_{b\in\mathcal{B}} P\big(\big\{\mathcal{E}_b > \delta_b(n\wedge m)^{1/2}/\big(3(x_{(n\wedge m)})^{1/2}\big)\big\} \cap B_\varepsilon(b_1)\big)\right) + \mathrm{o}(1) \\ &\quad\lesssim \left(\sum_{b\in\mathcal{B}} \frac{x_{(n\wedge m)}}{\delta_b^2(n\wedge m)(mb^{2\beta+2}\wedge 1)}\right) + \mathrm{o}(1),\end{aligned}$$

where we have used (20) in the last estimate. □

5.1.4. *Proof of Proposition 2.6*

Without loss of generality, we set $q_\tau = 0$. Recall definition (7) of the pseudo-estimator $\widehat{f_b}$ which knows the error distribution. We estimate

$$\begin{aligned}\sup_{x\in(-\zeta,\zeta)}\big|\widetilde{f_b}(x)-f(x)\big| &\le \sup_{x\in(-\zeta,\zeta)}\big|\widehat{f_b}(x)-f(x)\big| + \|\widetilde{f_b}-\widehat{f_b}\|_\infty \\ &\le \sup_{x\in(-\zeta,\zeta)}\big|\widehat{f_b}(x)-f(x)\big| + \left\|\frac{\varphi_K(bu)\varphi_n(u)}{\varphi_\varepsilon(u)}\left(\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)}-1\right)\right\|_{L^1}.\end{aligned}$$

The analysis of the first term is very classical. However, we are not aware of any reference in the given setup. Both terms will be treated separately in the following two steps. All estimates will be uniform in $f \in \mathcal{C}^\alpha(R,r,\zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R,\gamma)$.

Step 1: Let $b \in (0,1)$. We will show that there are constants $d, D > 0$ such that for any $t > d(b^\alpha + (nb^{2\beta+1})^{-2})$

$$P\Big(\sup_{x\in(-\zeta,\zeta)}\big|\widehat{f_b}(x)-f(x)\big| > t\Big) \le 2\exp\big(2\log n - Dnb^{(2\beta+1)}\big(t\wedge t^2\big)\big). \tag{27}$$

Then the result follows by choosing $t \sim b^\alpha + (\frac{\log n}{nb^{2\beta+1}})^{1/2}$. Let us define $x_k := -\zeta + kn^{-2}$ for $k = 1,\dots,\lfloor 2\zeta n^2\rfloor =: M$ as well as

$$\begin{aligned}\chi_j(x) &:= \mathcal{F}^{-1}\left[\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)}\mathrm{e}^{\mathrm{i}uY_j}\right](x) - \mathbb{E}\left[\mathcal{F}^{-1}\left[\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)}\mathrm{e}^{\mathrm{i}uY_j}\right](x)\right] \\ &= K_b * \mathcal{F}^{-1}\left[\mathbb{1}_{[-b^{-1},b^{-1}]}(u)\frac{\mathrm{e}^{\mathrm{i}uY_j}}{\varphi_\varepsilon(u)}\right](x) - K_b * f(x), \qquad x\in\mathbb{R}.\end{aligned}$$

Therefore, $\widehat{f_b}(x) - \mathbb{E}[\widehat{f_b}(x)] = \frac{1}{n}\sum_{j=1}^n \chi_j(x)$ and thus

$$\begin{aligned}\sup_{|x|<\zeta}\big|\widehat{f_b}(x)-f(x)\big| &\le \sup_{|x|<\zeta}\big|\mathbb{E}\big[\widehat{f_b}(x)\big]-f(x)\big| + \sup_{|x|<\zeta}\big|\widehat{f_b}(x)-\mathbb{E}\big[\widehat{f_b}(x)\big]\big| \\ &\le \sup_{|x|<\zeta}\big|\mathbb{E}\big[\widehat{f_b}(x)\big]-f(x)\big| + \sup_{|x|<\zeta}\min_{k=1,\dots,M}\left|\frac{1}{n}\sum_{j=1}^n\big(\chi_j(x)-\chi_j(x_k)\big)\right| \\ &\quad + \max_{k=1,\dots,M}\left|\frac{1}{n}\sum_{j=1}^n\chi_j(x_k)\right| \\ &=: B + V_1 + V_2.\end{aligned}$$

The bias term B can be bounded as in the classical density estimation setup (cf. also Fan [8], Thms. 1 and 2), noting that the constant does not depend on $x \in (-\zeta,\zeta)$. Hence, $|B| \lesssim b^\alpha$. Using a continuity argument and the properties of $f_\varepsilon \in \mathcal{D}^\beta(R,\gamma)$, the term V_1 can be bounded by

$$
\begin{aligned}
|V_1| &\le \frac{1}{n^2}\Big\|\frac{1}{n}\sum_{j=1}^{n}\chi_j'\Big\|_\infty \\
&= \frac{1}{n^3}\Big\|\sum_{j=1}^{n}\big(K_b'\big) * \Big(\mathcal{F}^{-1}\Big[\mathbb{1}_{[-b^{-1},b^{-1}]}(u)\frac{\mathrm{e}^{\mathrm{i}uY_j}}{\varphi_\varepsilon(u)}\Big] - f\Big)\Big\|_\infty \\
&\le \frac{1}{n^2 b}\|K'\|_{L^1}\big(\|\mathbb{1}_{[-b^{-1},b^{-1}]}\varphi_\varepsilon^{-1}\|_{L^1} + \|f\|_\infty\big) \lesssim n^{-2}b^{-(\beta+2)} \lesssim \big(nb^{2\beta+1}\big)^{-2}.
\end{aligned}
$$

Therefore, $|B + V_1| \le D_1(b^\alpha + (nb^{2\beta+1})^{-2})$ for some constant $D_1 > 0$. We obtain for all $t > d(b^\alpha + (nb^{2\beta+1})^{-2})$ with $d := 2D_1$

$$
\begin{aligned}
P\Big(\sup_{|x|<\zeta}\big|\widehat{f_b}(x) - f(x)\big| > t\Big) &\le P\Bigg(\max_{k=1,\dots,M}\Bigg|\frac{1}{n}\sum_{j=1}^{n}\chi_j(x_k)\Bigg| > \frac{t}{2}\Bigg) \\
&\le \sum_{k=1}^{M} P\Bigg(\Bigg|\frac{1}{n}\sum_{j=1}^{n}\chi_j(x_k)\Bigg| > \frac{t}{2}\Bigg).
\end{aligned}
$$

Finally, we will apply Bernstein's inequality. To this end, we estimate

$$
\max_{j,k}\big|\chi_j(x_k)\big| \le 2\|K_b\|_{L^1}\big\|\mathbb{1}_{[-b^{-1},b^{-1}]}\varphi_\varepsilon^{-1}\big\|_{L^1} \le D_2 b^{-(\beta+1)},
$$

with some constant $D_2 > 0$. Using Plancherel's identity, the variance can be estimated by

$$
\begin{aligned}
\operatorname{Var}\big(\chi_j(x_k)\big) &= \mathbb{E}\Bigg[\mathcal{F}^{-1}\Bigg[\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)}\mathrm{e}^{\mathrm{i}uY_j}\Bigg]^2(x_k)\Bigg] - (K_b * f)^2(x_k) \\
&\le \frac{1}{2\pi}\|f\|_\infty\Bigg\|\frac{\varphi_K(-bu)}{\varphi_\varepsilon(-u)}\Bigg\|_{L^2}^2 \lesssim D_3 b^{-(2\beta+1)},
\end{aligned}
$$

for some $D_3 > 0$. Then Bernstein's inequality yields

$$
\begin{aligned}
P\Big(\sup_{x\in(-\zeta,\zeta)}\big|\widehat{f_b}(x) - f(x)\big| > t\Big) &\le \sum_{k=1}^{M} P\Bigg(\Bigg|\sum_{j=1}^{n}\chi_j(x_k)\Bigg| > nt/2\Bigg) \\
&\le 2\exp\Bigg(\log M - \frac{nb^{(2\beta+1)}t^2}{8(D_3 + D_2 t/3)}\Bigg) \\
&\le 2\exp\big(2\log n - Dnb^{(2\beta+1)}\big(t \wedge t^2\big)\big),
\end{aligned}
$$

with some constant $D > 0$.

Step 2: By the Cauchy–Schwarz inequality, we have

$$
\begin{aligned}
&\mathbb{E}\biggl[\biggl\|\frac{\varphi_K(bu)\varphi_n(u)}{\varphi_\varepsilon(u)}\biggl(\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)}-1\biggr)\biggr\|_{L^1}\mathbb{1}_{B_\varepsilon(b)}\biggr] \\
&\quad\lesssim\biggl(\mathbb{E}\biggl[\biggl\|\frac{\varphi_n(u)}{\varphi_\varepsilon(u)}\mathbb{1}_{[-1/b,1/b]}(u)\biggr\|_{L^2}^2\biggr]\mathbb{E}\biggl[\biggl\|\biggl(\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)}-1\biggr)\mathbb{1}_{[-1/b,1/b]}(u)\biggr\|_{L^2}^2\mathbb{1}_{B_\varepsilon(b)}\biggr]\biggr)^{1/2} \\
&\quad\le\biggl(\|\varphi_X\|_{L^2}+\biggl(\int_{-1/b}^{1/b}\frac{\mathbb{E}[|\varphi_n(u)-\varphi_Y(u)|^2]}{|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u\biggr)^{1/2}\biggr) \\
&\qquad\times\biggl(\int_{-1/b}^{1/b}\mathbb{E}\biggl[\biggl|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)}-1\biggr|^2\mathbb{1}_{B_\varepsilon(b)}\biggr]\,\mathrm{d}u\biggr)^{1/2} \\
&\quad\lesssim\biggl(\|\varphi_X\|_{L^2}+\biggl(\frac{1}{nb^{2\beta+1}}\biggr)^{1/2}\biggr)\biggl(\frac{1}{mb^{2\beta+1}}\biggr)^{1/2},
\end{aligned}
$$

where we have used (25) for the last step. Therefore, the additional error due to the unknown error distribution satisfies for any $\delta>0$ by Markov's inequality and by Lemma 5.1

$$
\begin{aligned}
&P\biggl(\biggl\|\frac{\varphi_K(bu)\varphi_n(u)}{\varphi_\varepsilon(u)}\biggl(\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)}-1\biggr)\biggr\|_{L^1}>\delta\biggr) \\
&\quad\le\frac{1}{\delta}\mathbb{E}\biggl[\biggl\|\frac{\varphi_K(bu)\varphi_n(u)}{\varphi_\varepsilon(u)}\biggl(\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)}-1\biggr)\biggr\|_{L^1}\mathbb{1}_{B_\varepsilon(b)}\biggr]+P\Bigl(\inf_{|u|\le1/b}\bigl|\varphi_{\varepsilon,m}(u)\bigr|<m^{-1/2}\Bigr) \qquad (28)\\
&\quad\lesssim\frac{1}{\delta}\biggl(\frac{1}{mb^{2\beta+1}}\biggr)^{1/2}+\mathrm{o}(1)
\end{aligned}
$$

and thus $\|\widetilde{f}_b-\widehat{f}_b\|_\infty=\mathcal{O}_P((mb^{2\beta+1})^{-1/2})$. Note that the second term does not depend on δ and thus $\mathrm{o}(1)$ is sufficient.

5.1.5. *Proof of Theorems 2.7 and 2.8*

We start with a lemma that establishes consistency of the quantile estimator and then prove the theorems. To apply this lemma also for the adaptive result, we prove convergence uniformly over a set of bandwidths.

Lemma 5.4. *Grant Assumption A with $\ell=1$. Let $\mathcal{B}$ be a set of bandwidths satisfying $|\mathcal{B}|\lesssim\log n$, $\max\mathcal{B}\to0$ and $\min_{b\in\mathcal{B}}(\log n)^2/((n\wedge m)b^{2\beta+1})\to0$. Then*

$$
\sup_{f\in\mathcal{C}^\alpha(R,r,\zeta,U_n)}\ \sup_{f_\varepsilon\in\mathcal{D}^\beta(R,\gamma)}P\Bigl(\sup_{b\in\mathcal{B}}|\widetilde{q}_{\tau,b}-q_\tau|>\delta\Bigr)\to0\qquad\textit{for all }\delta>0.
$$

Proof. We follow the general strategy of the proof of Theorem 5.7 by van der Vaart [31] in the classical M-estimation setting. Recall the definition of $\widetilde{M}_b$ given in (2) and its deterministic counterpart $M(\eta) = \int_{-\infty}^{\eta} f(x)\,\mathrm{d}x - \tau$. To this end, we first claim that

$$\sup_{b\in\mathcal{B}} \widetilde{M}_b(\widetilde{q}_{\tau,b}) = \mathrm{o}_P(1), \tag{29}$$

Since $\widetilde{q}_{\tau,b}$ minimizes $\widetilde{M}_b$ on the interval $[-U_n, U_n]$ for $U_n \lesssim \log n$ and $M(q_\tau) = 0$ with $q_\tau \in [-U_n, U_n]$, Corollary 5.3 implies for any $\delta > 0$

$$\begin{aligned} P\Big(\sup_{b\in\mathcal{B}}\big|\widetilde{M}_b(\widetilde{q}_{\tau,b})\big| > \delta\Big) &\le P\Big(\sup_{b\in\mathcal{B}}\big|\widetilde{M}_b(q_\tau) - M(q_\tau)\big| > \delta\Big) \\ &= P\bigg(\sup_{b\in\mathcal{B}}\bigg|\int_{-\infty}^{q_\tau}\big(\widetilde{f}_b(x) - f(x)\big)\,\mathrm{d}x\bigg| > \delta\bigg) \to 0, \end{aligned} \tag{30}$$

which gives (29).

Now, we show that f satisfies the uniqueness condition

$$\inf_{\eta\,:\,|\eta-q_\tau|\ge\delta}\big|M(\eta)\big| > 0 \qquad \text{for any } \delta > 0. \tag{31}$$

By the Hölder regularity $M'(\eta) = f(\eta) \ge f(q_\tau) - |f(q_\tau) - f(\eta)| \ge r - R|q_\tau - \eta|^{1\wedge\alpha} \ge r/2$ for $|q_\tau - \eta| \le (\frac{r}{2R})^{1\vee\alpha^{-1}}$. Without loss of generality, we can assume $\delta \le (\frac{r}{2R})^{1\vee\alpha^{-1}}$, otherwise consider $\delta \wedge (\frac{r}{2R})^{1\vee\alpha^{-1}}$. Recall that q_τ is given by the root of M and that M is increasing. Hence, we obtain

$$\inf_{\eta\,:\,|\eta-q_\tau|\ge\delta}\big|M(\eta)\big| = \inf_{\eta\in\{-\delta,\delta\}}\big|M(q_\tau - \eta) - M(q_\tau)\big| \ge \delta \inf_{\eta\,:\,|\eta-q_\tau|\ge\delta} M'(\eta) \ge \frac{\delta r}{2}.$$

Applying (29) and (31) yield

$$\begin{aligned} P\Big(\sup_{b\in\mathcal{B}}|\widetilde{q}_{\tau,b} - q_\tau| > \delta\Big) &\le P\Big(\sup_{b\in\mathcal{B}}\big|M(\widetilde{q}_{\tau,b})\big| \ge \delta r/2\Big) \\ &= P\Big(\sup_{b\in\mathcal{B}}\big|M(\widetilde{q}_{\tau,b}) - \widetilde{M}_b(\widetilde{q}_{\tau,b})\big| \ge \delta r/3\Big) + \mathrm{o}(1) \\ &\le P\Big(\sup_{b\in\mathcal{B}}\sup_{\eta\in[-U_n,U_n]}\big|M(\eta) - \widetilde{M}_b(\eta)\big| \ge \delta r/3\Big) + \mathrm{o}(1) \\ &= P\bigg(\sup_{b\in\mathcal{B}}\sup_{\eta\in[-U_n,U_n]}\bigg|\int_{-\infty}^{\eta}\big(\widetilde{f}_b(x) - f(x)\big)\,\mathrm{d}x\bigg| \ge \delta r/3\bigg) + \mathrm{o}(1). \end{aligned} \tag{32}$$

Hence, it remains to show uniform consistency of $\int_{-\infty}^{\eta}\widetilde{f}_b(x)\,\mathrm{d}x$. Write

$$\begin{aligned} \bigg|\int_{-\infty}^{\eta}\big(\widetilde{f}_b(x) - f(x)\big)\,\mathrm{d}x\bigg| &\le \bigg|\int_{-\infty}^{\eta}\big(K_b * f(x) - f(x)\big)\,\mathrm{d}x\bigg| + \bigg|\int_{-\infty}^{\eta}\big(\widetilde{f}_b(x) - K_b * f(x)\big)\,\mathrm{d}x\bigg| \\ &= \big|K_b * F(\eta) - F(\eta)\big| + \bigg|\int_{-\infty}^{\eta}\big(\widetilde{f}_b(x) - K_b * f(x)\big)\,\mathrm{d}x\bigg|. \end{aligned}$$

We have $|K_b * F(\eta) - F(\eta)| = |\int K_b(z)(F(\eta - z) - F(\eta))\,\mathrm{d}z| \le b\|f\|_\infty \|zK(z)\|_{L^1}$ by the boundedness of f. Further note for $\eta \in [-U_n, U_n]$

$$\begin{aligned}
&\left|\int_{-\infty}^{\eta} \big(\widetilde{f}_b(x) - K_b * f(x)\big)\,\mathrm{d}x\right| \\
&\quad\le \left|\int_{-\infty}^{q_\tau} \big(\widetilde{f}_b(x) - K_b * f(x)\big)\,\mathrm{d}x\right| + \left|\int_{q_\tau\wedge\eta}^{q_\tau\vee\eta} \big(\widetilde{f}_b(x) - K_b * f(x)\big)\,\mathrm{d}x\right| \\
&\quad\le \left|\int_{-\infty}^{q_\tau} \big(\widetilde{f}_b(x) - K_b * f(x)\big)\,\mathrm{d}x\right| + \sqrt{2U_n}\left(\int_{-\infty}^{\infty} \big(\widetilde{f}_b(x) - K_b * f(x)\big)^2\,\mathrm{d}x\right)^{1/2},
\end{aligned}$$

where we have used the Cauchy–Schwarz inequality for the last step. Hence, together with (32) we obtain for all $\delta > 6\|f\|_\infty \|zK(z)\|_{L^1}/r \sup_{b\in\mathcal{B}} b$

$$\begin{aligned}
&P\Big(\sup_{b\in\mathcal{B}} |\widetilde{q}_{\tau,b} - q_\tau| > \delta\Big) \\
&\quad\le P\left(\sup_{b\in\mathcal{B}} \sup_{\eta\in[-U_n,U_n]} \left|\int_{-\infty}^{\eta} \big(\widetilde{f}_b(x) - f(x)\big)\,\mathrm{d}x\right| \ge \delta r/3\right) + \mathrm{o}(1) \\
&\quad\le P\left(\sup_{b\in\mathcal{B}} \left|\int_{-\infty}^{q_\tau} \big(\widetilde{f}_b(x) - K_b * f(x)\big)\,\mathrm{d}x\right| \ge \frac{\delta r}{9}\right) \\
&\qquad + P\left(\sup_{b\in\mathcal{B}} \int_{\mathbb{R}} \big(\widetilde{f}_b(x) - K_b * f(x)\big)^2\,\mathrm{d}x \ge \frac{\delta^2 r^2}{162 U_n}\right).
\end{aligned}$$

Corollary 5.3 shows under the conditions on $\mathcal{B}$ that

$$P\left(\sup_{b\in\mathcal{B}} \left|\int_{-\infty}^{q_\tau} \big(\widetilde{f}_b(x) - K_b * f(x)\big)\,\mathrm{d}x\right| > \delta r/9\right) \to 0.$$

Hence, it remains to show

$$P\left(\sup_{b\in\mathcal{B}} \int_{\mathbb{R}} \big(\widetilde{f}_b(x) - K_b * f(x)\big)^2\,\mathrm{d}x > \delta^2 r^2/(162 U_n)\right) \to 0. \tag{33}$$

On the event $B_\varepsilon(b)$, (33) follows basically from the work of Neumann [21]. More precisely, Plancherel's equality, (25) and the Cauchy–Schwarz inequality yield for any $b \in \mathcal{B}$

$$\begin{aligned}
&\mathbb{E}\left[\int_{\mathbb{R}} \big(\widetilde{f}_b(x) - K_b * f(x)\big)^2\,\mathrm{d}x \mathbb{1}_{B_\varepsilon(b)}\right] \\
&\quad= \frac{1}{2\pi} \int_{\mathbb{R}} \big|\varphi_K(bu)\big|^2 \mathbb{E}\left[\left|\frac{\varphi_n(u)}{\varphi_{\varepsilon,m}(u)} - \frac{\varphi_Y(u)}{\varphi_\varepsilon(u)}\right|^2 \mathbb{1}_{B_\varepsilon(b)}\right]\mathrm{d}u \\
&\quad\lesssim \int_{-1/b}^{1/b} \left(\mathbb{E}\left[\frac{|\varphi_n(u) - \varphi_Y(u)|^2}{|\varphi_{\varepsilon,m}(u)|^2} \mathbb{1}_{B_\varepsilon(b)}\right] + \big|\varphi_Y(u)\big|^2 \mathbb{E}\left[\left|\frac{1}{\varphi_{\varepsilon,m}(u)} - \frac{1}{\varphi_\varepsilon(u)}\right|^2 \mathbb{1}_{B_\varepsilon(b)}\right]\right)\mathrm{d}u
\end{aligned}$$

$$\lesssim \int_{-1/b}^{1/b} \left(\mathbb{E}\left[\frac{|\varphi_n(u) - \varphi_Y(u)|^2}{|\varphi_\varepsilon(u)|^2} \left(1 + m\left|\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)\right|^2\right)\right] + \frac{|\varphi_Y(u)|^2}{m|\varphi_\varepsilon(u)|^4} \right) \mathrm{d}u$$

$$\le \int_{-1/b}^{1/b} \frac{1}{|\varphi_\varepsilon(u)|^2}$$

$$\times \left(\left(\mathbb{E}\left[\left|\varphi_n(u) - \varphi_Y(u)\right|^4\right]\mathbb{E}\left[2 + 2m^2\left|\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)\right|^4\right]\right)^{1/2} + \frac{|\varphi_X(u)|^2}{m} \right) \mathrm{d}u$$

$$\lesssim \int_{-1/b}^{1/b} \left|\varphi_\varepsilon(u)\right|^{-2}\left(n^{-1} + m^{-1}\right) \mathrm{d}u \lesssim \frac{1}{(n \wedge m) b^{2\beta+1}}.$$

Using $B_\varepsilon(\min \mathcal{B}) \subseteq B_\varepsilon(b)$ and Lemma 5.1, (33) follows from Markov's inequality

$$P\left(\sup_{b\in\mathcal{B}} \int_{\mathbb{R}} \left(\widetilde{f}_b(x) - K_b * f(x)\right)^2 \mathrm{d}x > \delta^2 r^2/(162 U_n) \right)$$

$$\lesssim \frac{U_n}{\delta^2} \sum_{b\in\mathcal{B}} \mathbb{E}\left[\int_{\mathbb{R}} \left(\widetilde{f}_b(x) - K_b * f(x)\right)^2 \mathrm{d}x \mathbb{1}_{B_\varepsilon(\min\mathcal{B})} \right] + P\left(\left(B_\varepsilon(\min\mathcal{B})\right)^c\right)$$

$$\lesssim \frac{(\log n)^2}{\delta^2 (n \wedge m) b^{2\beta+1}} + \mathrm{o}(1). \qquad \square$$

Proof of Theorem 2.7. A Taylor expansion yields

$$\begin{aligned} \widetilde{q}_{\tau,b} - q_\tau &= \frac{\widetilde{M}_b(\widetilde{q}_{\tau,b}) - \widetilde{M}_b(q_\tau)}{\widetilde{M}_b'(q_\tau^*)} = \frac{\widetilde{M}_b(\widetilde{q}_{\tau,b}) - \int_{-\infty}^{q_\tau} \widetilde{f}_b(x)\,\mathrm{d}x + \tau}{\widetilde{f}_b(q_\tau^*)} \\ &= \frac{\widetilde{M}_b(\widetilde{q}_{\tau,b}) - \int_{-\infty}^{q_\tau} (\widetilde{f}_b(x) - f(x))\,\mathrm{d}x}{\widetilde{f}_b(q_\tau^*)}, \end{aligned} \tag{34}$$

for some intermediate point q_τ^* between q_τ and $\widetilde{q}_{\tau,b}$. By Proposition 2.5 and (30), the numerator in the above display is of order $\mathcal{O}_P(n^{-(\alpha+1)/(2\alpha+2\beta+1)})$ for the optimal bandwidth b^*. For the denominator, we will show $\widetilde{f}_b(q_\tau^*) = f(q_\tau) + \mathrm{o}_P(1)$ which completes the proof. Since $f(\cdot + q_\tau) \in C^\alpha([-\zeta, \zeta], R)$, we obtain $|f(x + q_\tau) - f(q_\tau)| < t/2$ for all $|x| \le (\frac{t}{2R})^{1\vee\alpha^{-1}} \wedge \zeta =: \delta$ for any $t > 0$. Therefore,

$$\begin{aligned} &P\left(\left|\widetilde{f}_b(q_\tau^*) - f(q_\tau)\right| > t\right) \\ &\quad\le P\Big(\sup_{x\in[-\delta,\delta]} \left|\widetilde{f}_b(x + q_\tau) - f(q_\tau)\right| > t\Big) + P\left(|\widetilde{q}_{\tau,b} - q_\tau| > \delta\right) \\ &\quad\le P\Big(\sup_{x\in[-\delta,\delta]} \left|\widetilde{f}_b(x + q_\tau) - f(x + q_\tau)\right| > t/2\Big) + P\left(|\widetilde{q}_{\tau,b} - q_\tau| > \delta\right). \end{aligned} \tag{35}$$

Checking that the bandwidth satisfies $b \to 0$ and $\log(n)/(nb^{2\beta+1}) \to 0$ for $n \to \infty$, the first term on the right-hand side above converges to zero by the uniform consistency proved in Proposition 2.6. The second one vanishes asymptotically by Lemma 5.4. $\square$

Proof of Theorem 2.8. Under the smoothness condition the interval (τ_1, τ_2) coincides with a bounded interval of quantiles (q_{τ_1}, q_{τ_2}). Noting that all our estimates are independent of the quantile, Theorem 2.8 can be proved along the same lines as Theorem 2.7 with only minor adaptation to $\sup_{\tau\in(\tau_1,\tau_2)}$ given a uniform version of Proposition 2.5: uniformly over f in the class defined in the theorem and $f_\varepsilon \in \mathcal{D}^\beta(R,\gamma)$ for any b such that $(n \wedge m)b^{2\beta+1} \to \infty$ it holds

$$\begin{aligned}
&\sup_{\tau\in(\tau_1,\tau_2)}\left|\int_{-\infty}^{q_\tau}\big(\widetilde{f}_b(x)-f(x)\big)\,\mathrm{d}x\right| \\
&\quad= \mathcal{O}_P\left(b^{\alpha+1} + \left(\frac{\log n}{n}\vee\frac{1}{m}\right)^{1/2}\big(b^{-\beta+1/2}\vee 1\big) + \left(\frac{1}{n}\vee\frac{1}{m}\right)^{1/2}\big(m^{-1/2}b^{-\beta-1}\vee 1\big)\right).
\end{aligned}\tag{36}$$

Hence, when $\alpha \geq 1/2$ the asymptotically optimal choice $b = (\frac{\log n}{n} \wedge \frac{1}{m})^{1/(2\alpha+2(\beta\vee 1/2)+1)}$ yields

$$\left|\int_{-\infty}^{q_\tau}\big(\widetilde{f}_b(x)-f(x)\big)\,\mathrm{d}x\right| = \mathcal{O}_P\left(\left(\frac{\log n}{n}\vee\frac{1}{m}\right)^{(\alpha+1)/(2\alpha+2\beta+1)} \vee \left(\frac{\log n}{n}\vee\frac{1}{m}\right)^{1/2}\right).$$

The result (36) can be obtained as Proposition 2.5 except for the term $T_{s,x} = T_{s,x}(q_\tau)$, defined in (22), which will be treated in the following. Defining the grid $\tau_1 = \sigma_0 \leq \cdots \leq \sigma_M = \tau_2$ such that $q_{\sigma_{k+1}} - q_{\sigma_k} \leq (q_{\tau_2} - q_{\tau_1})/M$ for $k = 1, \ldots, M$ and $M \in \mathbb{N}$, we decompose for any $c > 0$

$$\begin{aligned}
P\Big(\sup_{\tau\in(\tau_1,\tau_2)}\big|T_{s,x}(q_\tau)\big| > c\Big) &\leq P\Big(\max_{k=1,\ldots,M}\big|T_{s,x}(q_{\sigma_k})\big| > c/2\Big) \\
&\quad + P\Big(\sup_{\substack{q_1,q_2\in(q_{\tau_1},q_{\tau_2}):\\ |q_1-q_2|\leq(q_{\tau_2}-q_{\tau_1})/(2M)}}\big|T_{s,x}(q_1)-T_{s,x}(q_2)\big| > c/2\Big).
\end{aligned}\tag{37}$$

For the first term, we deduce a concentration inequality. We write

$$\frac{1}{2\pi}T_{s,x} = \frac{1}{2\pi}T_{s,x}(q_\tau) = \frac{1}{n}\sum_{j=1}^{n}\big(\xi_{j,b}(q_\tau) - \mathbb{E}\big[\xi_{j,b}(q_\tau)\big]\big)$$

with

$$\xi_{j,b}(q_\tau) = \int_{-\infty}^{0} a_s(x)\mathcal{F}^{-1}\left[\frac{\varphi_K(bu)\mathrm{e}^{\mathrm{i}uY_j}}{\varphi_\varepsilon(u)}\right](x+q_\tau)\,\mathrm{d}x = \mathcal{F}^{-1}\left[\mathcal{F}a_s(-u)\frac{\varphi_K(bu)\mathrm{e}^{\mathrm{i}uY_j}}{\varphi_\varepsilon(u)}\right](q_\tau).$$

Uniformly in q_τ we have the deterministic bound

$$\big|\xi_{j,b}(q_\tau)\big| \leq \frac{1}{2\pi}\int_{-1/b}^{1/b}\big|\mathcal{F}a_s(-u)\big|\left|\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)}\right|\mathrm{d}u \lesssim \int_{-1/b}^{1/b}\frac{1}{(1+|u|)|\varphi_\varepsilon(u)|}\,\mathrm{d}u \lesssim b^{-\beta}. \tag{38}$$

Hence, $|\xi_{j,b}(q_\tau) - \mathbb{E}[\xi_{j,b}(q_\tau)]| \lesssim b^{-\beta}$. Since the variance of $T_{s,x}(q_\tau)$ is bounded by (23), Bernstein's inequality (e.g., Massart [19], Prop. 2.9) yields for some constant $C > 0$ independent of q_τ

$$P\big(\big|T_{s,x}(q_\tau)\big| \geq \kappa\big(n^{-1/2}b^{-\beta+1/2} \vee n^{-1/2}\big)\big) \leq 2\exp\bigg(-\frac{C\kappa^2}{1+\kappa(nb)^{-1/2}}\bigg).$$

For the second term on the right-hand side of (37), we estimate

$$\begin{aligned}\big|T_{s,x}(q_1) - T_{s,x}(q_2)\big| &\leq \bigg\|\bigg(\mathcal{F}^{-1}\bigg[\mathcal{F}a_s(u)\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)}\big(\varphi_n(u) - \varphi_Y(u)\big)\bigg]\bigg)'\bigg\|_\infty |q_1 - q_2| \\ &\leq \frac{|q_1 - q_2|}{2\pi}\int_{\mathbb{R}} |u|\big|\mathcal{F}a_s(u)\big|\frac{|\varphi_K(bu)|}{|\varphi_\varepsilon(u)|}\big|\varphi_n(u) - \varphi_Y(u)\big|\,\mathrm{d}u \\ &\lesssim |q_1 - q_2|\int_{-1/b}^{1/b}\big(1+|u|\big)^\beta\big|\varphi_n(u) - \varphi_Y(u)\big|\,\mathrm{d}u.\end{aligned}$$

Using Markov's inequality, we thus estimate (37) by

$$\begin{aligned}&P\Big(\sup_{\tau\in(\tau_1,\tau_2)}\big|T_{s,x}(q_\tau)\big| > \kappa\big(n^{-1/2}b^{-\beta+1/2} \vee n^{-1/2}\big)\Big) \\ &\quad\lesssim M\exp\bigg(-\frac{C\kappa^2}{4+2\kappa(nb)^{-1/2}}\bigg) \\ &\qquad + \frac{(q_{\tau_2} - q_{\tau_1})n^{1/2}(b^{\beta-1/2}\wedge 1)}{M\kappa}\mathbb{E}\bigg[\int_{-1/b}^{1/b}\big(1+|u|\big)^\beta\big|\varphi_n(u) - \varphi_Y(u)\big|\,\mathrm{d}u\bigg] \\ &\quad\lesssim M\exp\bigg(-\frac{C\kappa^2}{4+2\kappa(nb)^{-1/2}}\bigg) + \frac{(q_{\tau_2} - q_{\tau_1})(b^{-3/2}\wedge(b^{-\beta-1}))}{M\kappa}.\end{aligned}$$

Choosing $M = n^2$ and $\kappa = (\frac{9}{C}\log n)^{1/2}$, we have $\kappa(nb)^{-1/2} = \mathrm{o}(1)$ and the previous display converges to zero. Hence,

$$\sup_{\tau\in(\tau_1,\tau_2)}\big|T_{s,x}(q_\tau)\big| = \mathcal{O}_P\bigg(\bigg(\frac{\log n}{n}\bigg)^{1/2}b^{-\beta+1/2} \vee \bigg(\frac{\log n}{n}\bigg)^{1/2}\bigg). \qquad \square$$

5.1.6. *Proof of Theorem* 2.10

To prove the lower bound for the estimation of the distribution function, we can assume without loss of generality $q = 0$. For $n \leq m$ the estimation error of $\bar{F}_{n,m}(0)$ is bounded from below by the estimation error with known error distribution. A lower bound for the latter is proved by Fan [8] whose construction can be used in our setting, too.

To prove the lower bound for $m < n$, we will apply Theorem 2.1 in Tsybakov [30]. To this end, we construct two alternatives $(F_i, f_{\varepsilon,i}) \in \widetilde{\mathcal{C}}^{\alpha+1}(R, r, [-\zeta,\zeta]) \times \mathcal{D}^\beta(R,\gamma)$, $i = 1, 2$, such that the χ^2-distance of the corresponding laws of $(Y_1,\dots,Y_n,\varepsilon_1^*,\dots,\varepsilon_m^*)$ is bounded by some small constant and such that $|F_1(0) - F_2(0)|$ is bounded from below with the right rate. Recall that the convolution of a c.d.f. F with a function g is defined as $F * g(x) = \int g(x-y)\,\mathrm{d}F(y)$. Following the idea by Neumann [21] our construction will satisfy $F_1 * f_{\varepsilon,1} = F_2 * f_{\varepsilon,2}$ and is thus independent of n.

Step 1: For the construction of the alternatives, we need the following: let f_0 be a bounded density whose corresponding distribution is in $\mathcal{C}^{\alpha+1}(R, r, \zeta)$ satisfying $q_\tau = 0$. Let $f_{\varepsilon,0}$ be an inner point of $\mathcal{D}^\beta(R, \gamma)$ with

$$f_{\varepsilon,0}(x) \gtrsim \bigl(1+|x|\bigr)^{-\gamma-2}, \qquad \bigl|(\mathcal{F} f_{\varepsilon,0})^{(k)}(u)\bigr| \lesssim \bigl(1+|u|\bigr)^{-\beta}, \qquad k = 0, \dots, K \tag{39}$$

for $x, u \in \mathbb{R}$ and an integer $K > \gamma/2 + 1$. Let the perturbation $g \in C^\infty(\mathbb{R}) \cap L^1(\mathbb{R})$ satisfy

$$\int_{\mathbb{R}} g(x)\,\mathrm{d}x = 0, \qquad \int_{-\infty}^0 g(x)\,\mathrm{d}x \neq 0,$$
$$\bigl\|\bigl(1 \vee x^{\gamma\vee 1}\bigr) g(x)\bigr\|_{L^1} < \infty, \qquad \operatorname{supp} \mathcal{F}g \subseteq [-2,-1] \cup [1,2].$$

Define $g_b := b^{-1} g(\cdot/b)$ for $b > 0$ and for some $a \in (0,1)$, $c > 0$

$$\begin{aligned} F_1(x) &:= a \int_{-\infty}^x f_0(y)\,\mathrm{d}y + (1-a)\mathbb{1}_{[2\zeta,\infty)}(x), \\ f_{\varepsilon,1}(x) &:= f_{\varepsilon,0} + c b^{\alpha+1}\bigl(f_{\varepsilon,0} * g_b(\cdot + 2\zeta)\bigr)(x), \\ F_2(x) &:= F_1(x) + c b^{\alpha+1} \int_{-\infty}^x g_b(\cdot + 2\zeta) * F_1(y)\,\mathrm{d}y, \\ f_{\varepsilon,2}(x) &:= f_{\varepsilon,0}(x). \end{aligned} \tag{40}$$

Owing to $\int g_b = 0$, F_i are distribution functions admitting Lebesgue densities on $[-\zeta, \zeta]$ which are at least α-Hölder continuous. Estimating $\|f_0 * g_b\|_{C^\alpha(\mathbb{R})} \lesssim \|f_0\|_{L^1} \|g_b\|_{C^\alpha(\mathbb{R})} \lesssim b^{-\alpha-1}$, we infer that $\mathrm{d}F_2$ is contained in a closed Hölder ball. Hence, $F_i \in \widetilde{\mathcal{C}}^{\alpha+1}(R, r, [-\zeta, \zeta])$ for $c > 0$ sufficiently small. $f_{\varepsilon_i} \in \mathcal{D}^\beta(R, \gamma)$ can be verified, using $\int g = 0$, $\|\mathcal{F}g\|_\infty \le \|g\|_{L^1}$ and $\|(\mathcal{F}g)'(u)(1+|u|)\|_\infty < \infty$.

Step 2: To bound the distance $|F_1(0) - F_2(0)|$ from below we note, using Fubini's theorem, $\int g = 0$ and $\|f_0\|_\infty < \infty$,

$$\begin{aligned} F_2(0) - F_1(0) &= b^{\alpha+1}\biggl(ac \int_{\mathbb{R}} \int_{2\zeta}^{-y+2\zeta} f_0(x) g_b(y)\,\mathrm{d}x\,\mathrm{d}y + (1-a)c \int_{-\infty}^0 g_b(x)\,\mathrm{d}x\biggr) \\ &= b^{\alpha+1}\biggl((1-a)c \int_{-\infty}^0 g(x)\,\mathrm{d}x + \mathcal{O}\bigl(\bigl\|y g_b(y)\bigr\|_{L^1}\bigr)\biggr) \\ &= b^{\alpha+1}\biggl((1-a)c \int_{-\infty}^0 g(x)\,\mathrm{d}x + \mathcal{O}(b)\biggr), \end{aligned} \tag{41}$$

for b small enough. Therefore, $|F_1(0) - F_2(0)| \gtrsim b^{\alpha+1}$.

Step 3: Using the independence of the observations, the sample $(Y_1, \dots, Y_n, \varepsilon_1^*, \dots, \varepsilon_m^*)$ is distributed according to $(F_i * f_{\varepsilon,i})^{\otimes n} \otimes f_{\varepsilon,i}^{\otimes m}$ under the hypotheses $i = 1, 2$. By construction $F_1 * f_{\varepsilon,1} = F_2 * f_{\varepsilon,2}$ such that the χ^2-distance of the laws of the observations equals

$$\chi^2\big(f_{\varepsilon,1}^{\otimes m}, f_{\varepsilon,2}^{\otimes m}\big) = \bigg(1 + \int_{\mathbb{R}} \frac{(f_{\varepsilon,1} - f_{\varepsilon,2})^2(x)}{f_{\varepsilon,2}(x)} \,\mathrm{d}x\bigg)^m - 1. \tag{42}$$

We decompose

$$\begin{aligned}
&\int_{\mathbb{R}} \frac{(f_{\varepsilon,1} - f_{\varepsilon,2})^2(x)}{f_{\varepsilon,2}(x)} \,\mathrm{d}x \\
&\quad = c^2 b^{2\alpha+2}\bigg(\int_{|x|\le 1} \frac{(f_{\varepsilon,0} * g_b(\cdot + 2\zeta))^2(x)}{f_{\varepsilon,0}(x)} \,\mathrm{d}x + \int_{|x|>1} \frac{(f_{\varepsilon,0} * g_b(\cdot + 2\zeta))^2(x)}{f_{\varepsilon,0}(x)} \,\mathrm{d}x\bigg) \\
&\quad =: c^2 b^{2\alpha+2}(I_1 + I_2).
\end{aligned}$$

For the first integral, we use $\inf_{|x|\le 1} f_{\varepsilon,0}(x) > 0$, Plancherel's identity, $f_{\varepsilon,0} \in \mathcal{D}^\beta(R, \gamma)$ and the support of $\mathcal{F}g$ to estimate

$$|I_1| \lesssim \int_{\mathbb{R}} \big|\mathcal{F} f_{\varepsilon,0}(u)\mathcal{F}g(bu)\mathrm{e}^{-\mathrm{i}2\zeta u}\big|^2 \,\mathrm{d}u \lesssim \int_{1/b\le|u|\le 2/b} \big(1 + |u|\big)^{-2\beta} \,\mathrm{d}u \lesssim b^{2\beta-1}.$$

Using (39), I_2 can be estimated similarly

$$\begin{aligned}
|I_2| &\lesssim \int_{|x|>1} \big(1 + |x|\big)^{\gamma+2} |x|^{-2K} \big|\mathcal{F}^{-1}\big[\big(\mathcal{F} f_{\varepsilon,0}\mathcal{F}g_b \mathrm{e}^{-\mathrm{i}2\zeta\cdot}\big)^{(K)}\big]\big|^2(x) \,\mathrm{d}x \\
&\sim \int_{1/b\le|u|\le 2/b} \big|\big(\mathcal{F} f_{\varepsilon,0}(u)\mathcal{F}g(bu)\mathrm{e}^{-\mathrm{i}2\zeta u}\big)^{(K)}\big|^2 \,\mathrm{d}u \lesssim b^{2\beta-1}.
\end{aligned}$$

We conclude from (42) for some constant $C > 0$ that

$$\chi^2\big(f_{\varepsilon,1}^{\otimes m}, f_{\varepsilon,2}^{\otimes m}\big) \le \big(1 + Cc^2 b^{2\alpha+2\beta+1}\big)^m - 1 \le \exp\big(Cc^2 m b^{2\alpha+2\beta+1}\big) - 1,$$

which can be bounded by an arbitrarily small constant if c is chosen sufficiently small and $b = m^{-1/(2\alpha+2\beta+1)}$. We obtain from Step 2 that $|F_1(0) - F_2(0)| \ge Cm^{-(\alpha+1)/(2\alpha+2\beta+1)}$, for some positive constant C.

Step 4: Replacing in (40) the factor $b^{\alpha+1}$ in F_2 and $f_{\varepsilon,1}$ by $cm^{-1/2}$ for some sufficiently small constant $c > 0$ and choosing $b = 1$, the previous steps yield the lower bound $m^{-1/2}$.

Let us finally conclude the lower bound for the estimation error of the quantiles. We use the construction from Step 1, denoting the τ-quantile of F_i by $q_{\tau,i}$. We note $|q_{\tau,1}| < \delta$ for any $\delta > 0$ if we choose a close enough to one and thus F_1 is regular in an interval around $q_{\tau,1}$. Moreover, it holds

$$\begin{aligned}
\|F_1 - F_2\|_\infty &\le c\big(m^{-1/2} \vee b^{\alpha+1}\big)\big\|\big(af_0 + (1-a)\delta_{-2\zeta}\big) * g_b(\cdot + 2\zeta)\big\|_{L^1} \\
&\le c\big(m^{-1/2} \vee b^{\alpha+1}\big)\|g\|_{L^1} \to 0.
\end{aligned}$$

We infer analogously to (32) that $|q_{\tau,1} - q_{\tau,2}| < \delta$ for any $\delta > 0$ and m sufficiently large implying $F_i \in \widetilde{\mathcal{C}}^{\alpha+1}(R, r, \zeta)$. Applying a Taylor expansion similar to (4), we obtain

$$q_{\tau,2} - q_{\tau,1} = -\frac{F_2(q_{\tau,1}) - F_1(q_{\tau,1})}{F_2'(q_\tau^*)}$$

for some intermediate point between $q_{\tau,1}$ and $q_{\tau,2}$. The denominator $F_2'(q_\tau^*)$ is bounded from above and below owing to $\sup_{|x|\le\zeta} |F_2'(x) - af_0(x)| \to 0$, $|q_{\tau,2}| \le |q_{\tau,2} - q_{\tau,1}| + |q_{\tau,1}| < 2\delta$ and $f_0(0) > 0$. (41) yields $|q_{\tau,2} - q_{\tau,1}| \gtrsim m^{-1/2} \vee b^{\alpha+1}$. The assertion follows from Steps 3 and 4 above.

5.2. Proofs for Section 3

We start with Lemma 3.1 concerning the bandwidth set $\mathcal{B}_n$ from (9).

5.2.1. *Proof of Lemma* 3.1

By Lemma 5.1, we can argue on the event $B_\varepsilon(b)$ from (3). The deterministic counterpart of $\widetilde{j}_n$, defined in (8), is given by

$$j_{0,n} := \min\left\{ j = 0, \dots, N_n \colon 2 \le \left(\frac{\log n}{n}\right)^{1/2} \int_{-1/b_j}^{1/b_j} \left|\varphi_\varepsilon(u)\right|^{-1} \mathrm{d}u \le 4 \right\}. \tag{43}$$

Noting that for $f_\varepsilon \in \mathcal{D}^\beta(R, \gamma)$

$$4 \ge \left(\frac{\log n}{n}\right)^{1/2} \int_{-1/b_{j_{0,n}}}^{1/b_{j_{0,n}}} \left|\varphi_\varepsilon(u)\right|^{-1} \mathrm{d}u \gtrsim \left(\frac{\log n}{n b_{j_{0,n}}^{2\beta+2}}\right)^{1/2}$$

we obtain $n b_{j_{0,n}}^{2\beta+2} \to \infty$ and thus it is sufficient to prove

$$\inf_{f\in\mathcal{C}^\alpha(R,r,\zeta)} \inf_{f_\varepsilon\in\mathcal{D}^\beta(R,\gamma)} P\left(\left\{b_{j_{0,n}} \le b_{\widetilde{j}_n} \le b^*\right\} \cap B_\varepsilon(b_{j_{0,n}})\right) \to 1 \qquad \text{as } n \to \infty, \tag{44}$$

for the optimal bandwidth $b^* = n^{-1/(2\alpha+2(\beta\vee1/2)+1)}$. For convenience, we define

$$I_n(b) := \left(\frac{\log n}{n}\right)^{1/2} \int_{-1/b}^{1/b} \frac{\mathrm{d}u}{|\varphi_\varepsilon(u)|}, \qquad \widetilde{I}_n(b) := \left(\frac{\log n}{n}\right)^{1/2} \int_{-1/b}^{1/b} \frac{\mathrm{d}u}{|\varphi_{\varepsilon,m}(u)|}.$$

Assume $b_{\widetilde{j}_n} < b_{j_{0,n}}$, then monotonicity implies $\widetilde{I}_n(b_{j_{0,n}}) \le \widetilde{I}_n(b_{\widetilde{j}_n}) \le 1$. Combined with $I_n(j_{0,n}) \ge 2$, we obtain $I_n(b_{j_{0,n}}) - \widetilde{I}_n(b_{j_{0,n}}) \ge 1$. Hence,

$$\{b_{\widetilde{j}_n} < b_{j_{0,n}}\} \subseteq \left\{\left|I_n(b_{j_{0,n}}) - \widetilde{I}_n(b_{j_{0,n}})\right| \ge 1\right\}. \tag{45}$$

On the other hand, if $b^* < b_{\widetilde{j}_n}$, we get $\widetilde{I}_n(b^*) \geq \widetilde{I}_n(b_{\widetilde{j}_n}) \geq 1/2$. Since $I_n(b^*) \lesssim (\frac{\log n}{n(b^*)^{2\beta+2}})^{1/2}$ converges to zero, $I_n(b^*) \leq 1/4$ for n large enough. Thus,

$$\{b_{\widetilde{j}_n} > b^*\} \subseteq \{|I_n(b^*) - \widetilde{I}_n(b^*)| \geq 1/4\}. \tag{46}$$

To show that the probabilities of the right-hand sides of (45) and (46) converge to zero, we first apply the Cauchy–Schwarz inequality

$$\begin{aligned}
|I_n(b) - \widetilde{I}_n(b)|^2 &\leq \frac{\log n}{n} \int_{-1/b}^{1/b} \frac{\mathrm{d}u}{|\varphi_\varepsilon(u)|^2} \int_{-1/b}^{1/b} \left|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1\right|^2 \mathrm{d}u \\
&\lesssim \frac{\log n}{nb^{2\beta+1}} \int_{-1/b}^{1/b} \left|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1\right|^2 \mathrm{d}u.
\end{aligned}$$

Markov's inequality and (25) yield for $b \in \{b_{\min}, b^*\}$

$$\begin{aligned}
P\left(\left\{|I_n(b) - \widetilde{I}_n(b)| \geq \frac{1}{4}\right\} \cap B_\varepsilon(b_{j_{0,n}})\right) &\lesssim \frac{\log n}{nb^{2\beta+1}} \int_{-1/b}^{1/b} \mathbb{E}\left[\left|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1\right|^2 \mathbb{1}_{B_\varepsilon(b_{j_{0,n}})}\right] \mathrm{d}u \\
&\lesssim \frac{\log n}{nmb^{4\beta+2}}
\end{aligned}$$

which converges to zero. Therefore, (44) holds true.

5.2.2. *Preparations to the Proof of Theorem* 3.2

Before we can prove Theorem 3.2, some preparations are needed. By Lemma 5.2 there is a constant $D > 0$ such that the bias can be bounded by $B_b := Db^{\alpha+1}$. By the error representation (34), we have for any $b \in \mathcal{B}$

$$\begin{aligned}
|\widetilde{q}_{\tau,b} - q_\tau| &= \left|\frac{\int_{-\infty}^{q_\tau} (\widetilde{f}_b(x) - f(x))\,\mathrm{d}x - \widetilde{M}_b(\widetilde{q}_{\tau,b})}{\widetilde{f}_b(\widetilde{q}^*)}\right| \\
&\leq \frac{B_b + |V_{b,X} + V_{b,\varepsilon} + V_{b,c}| + |\widetilde{M}_b(\widetilde{q}_{\tau,b})|}{|\widetilde{f}_b(q^*)|}
\end{aligned} \tag{47}$$

with some $q^* \in [(q_\tau \wedge \widetilde{q}_{\tau,b}), (q_\tau \vee \widetilde{q}_{\tau,b})]$ and where the stochastic error is decomposed in

$$\begin{aligned}
V_{b,X} &:= \frac{1}{n} \sum_{j=1}^{n} (\xi_j(b) - \mathbb{E}[\xi_j(b)]) \qquad \text{with} \\
\xi_j(b) &:= \int_{-\infty}^{0} a_s(x) \mathcal{F}^{-1}\left[\frac{\varphi_K(bu)\mathrm{e}^{\mathrm{i}uY_j}}{\varphi_\varepsilon(u)}\right](x + q_\tau)\,\mathrm{d}x, \\
V_{b,\varepsilon} &:= \int_{-\infty}^{0} a_s(x) \mathcal{F}^{-1}\left[\frac{\varphi_K(bu)\varphi_n(u)}{\varphi_\varepsilon(u)}\left(\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1\right)\right](x + q_\tau)\,\mathrm{d}x,
\end{aligned} \tag{48}$$

$$V_{b,c} := \int_{-\infty}^{0} a_c(x)\mathcal{F}^{-1}\left[\varphi_K(bu)\left(\frac{\varphi_n(u)}{\varphi_{\varepsilon,m}(u)} - \varphi_X(u)\right)\right](x+q_\tau)\,\mathrm{d}x.$$

In view of the analysis in Section 5.1.3, the part of the stochastic error which is due to the continuous part a_c will be negligible. Hence, we concentrate on $V_{b,X}$ and $V_{b,\varepsilon}$. By independence of $(\xi_j(b))_j$, we obtain

$$\begin{aligned}\operatorname{Var}(V_{b,X}) &\le \frac{1}{n}\mathbb{E}\big[\xi_j(b)^2\big] = \frac{1}{n}\mathbb{E}\left[\left(\int_{-\infty}^{0} a_s(x)\mathcal{F}^{-1}\left[\frac{\varphi_K(bu)\mathrm{e}^{\mathrm{i}uY_j}}{\varphi_\varepsilon(u)}\right](x+q_\tau)\,\mathrm{d}x\right)^2\right] \\ &=: \sigma_{b,X}^2.\end{aligned} \tag{49}$$

We will determine the variance of $V_{b,\varepsilon}$ on the event $B_\varepsilon(b)$, defined in (3). We apply Plancherel's identity and the Cauchy–Schwarz inequality to separate Y_i and ε_i from each other:

$$\begin{aligned}&\mathbb{E}\big[|V_{b,\varepsilon}|\mathbb{1}_{B_\varepsilon(b)}\big] \\ &\quad= \frac{1}{2\pi}\mathbb{E}\left[\left|\int_{\mathbb{R}} \mathcal{F}a_s(-u)\mathrm{e}^{-\mathrm{i}uq_\tau}\frac{\varphi_K(bu)\varphi_n(u)}{\varphi_\varepsilon(u)}\left(\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1\right)\mathrm{d}u\right|\mathbb{1}_{B_\varepsilon(b)}\right] \\ &\quad\le \frac{1}{2\pi}\mathbb{E}\left[\left(\int_{\mathbb{R}}\big|\varphi_K(bu)\big|\left|\frac{\varphi_n(u)}{\varphi_\varepsilon(u)}\right|^2\mathrm{d}u\right)^{1/2}\right. \\ &\qquad\qquad \left.\times\left(\int_{\mathbb{R}}\big|\varphi_K(bu)\big|\big|\mathcal{F}a_s(-u)\big|^2\left|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)} - 1\right|^2\mathrm{d}u\right)^{1/2}\mathbb{1}_{B_\varepsilon(b)}\right] \\ &\quad\le \frac{1}{2\pi}\mathbb{E}\left[\left(\int_{\mathbb{R}}\big|\varphi_K(bu)\big|\left|\frac{\varphi_n(u)}{\varphi_\varepsilon(u)}\right|^2\mathrm{d}u\right)^{1/2}\left(\int_{\mathbb{R}}\big|\varphi_K(bu)\big|\left|\frac{\mathcal{F}a_s(-u)}{\varphi_{\varepsilon,m}(u)}\right|^2\mathrm{d}u\mathbb{1}_{B_\varepsilon(b)}\right)^{1/2}\right. \\ &\qquad\qquad \left.\times\sup_{|u|\le 1/b}\big|\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)\big|\right].\end{aligned} \tag{50}$$

Let us define

$$\sigma_{b,\varepsilon} := \frac{1}{2\pi}m^{-1/2}\sigma_{b,\varepsilon,1}\sigma_{b,\varepsilon,2} \tag{51}$$

with

$$\begin{aligned}\sigma_{b,\varepsilon,1} &:= \mathbb{E}\left[\left(\int_{\mathbb{R}}\big|\varphi_K(bu)\big|\left|\frac{\varphi_n(u)}{\varphi_\varepsilon(u)}\right|^2\mathrm{d}u\right)^{1/2}\right], \\ \sigma_{b,\varepsilon,2} &:= \mathbb{E}\left[\left(\int_{\mathbb{R}}\big|\varphi_K(bu)\big|\left|\frac{\mathcal{F}a_s(-u)}{\varphi_{\varepsilon,m}(u)}\right|^2\mathrm{d}u\right)^{1/2}\mathbb{1}_{B_\varepsilon(b)}\right].\end{aligned}$$

With the bounds $\sigma_{b,X}$ and $\sigma_{b,\varepsilon}$ at hand, we obtain the following concentration results.

Lemma 5.5. *Let $\mathcal{B}$ be a set satisfying $|\mathcal{B}| \lesssim \log n$, $(\log\log n)/nb_1 \to 0$ for $b_1 = \min\mathcal{B}$ as well as $|\log b_1| \lesssim \log n$. Then we obtain uniformly over $f \in \mathcal{C}^\alpha(R, r, \zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R,\gamma)$ for any $\delta > 0$:*

(i) $P(\exists b \in \mathcal{B}\colon |V_{b,X}| \ge (1+\delta)\sqrt{\log\log n}(\sqrt{2}\sigma_{b,X} + \mathrm{o}(n^{-1/2}(b^{-\beta+1/2} \vee 1)))) \to 0.$

(ii) $P(\exists b \in \mathcal{B}\colon |V_{b,\varepsilon}| \ge \delta(\log n)^3\sigma_{b,\varepsilon}) \to 0.$

(iii) *Assuming further* $mb_1^{(2\beta\wedge 1)+2} \gtrsim 1$,
$P(\exists b \in \mathcal{B}\colon |V_{b,c}| \ge (\log n)^{3/2} n^{-1/2}(b^{-\beta+1/2} \vee 1)) \to 0.$

Proof. (i) Using the deterministic bound (38), we obtain $|\xi_j(b) - \mathbb{E}[\xi_j(b)]| \le Cb^{-\beta}$ for some constant $C > 0$. Since the variance is bounded by (49), Bernstein's inequality (e.g., Massart [19], Prop. 2.9) yields for any positive $\kappa_n = \mathrm{o}(nb)$

$$P\left(|V_{b,X}| \ge \sqrt{2\sigma_{b,X}^2\kappa_n} + \frac{C\kappa_n}{3nb^\beta}\right) \le 2\mathrm{e}^{-\kappa_n}.$$

Hence, $\sqrt{\kappa_n}(nb^\beta)^{-1} \lesssim (n(b^{2\beta-1} \wedge 1))^{-1/2}(\kappa_n/(nb))^{1/2}$ yields uniformly in $\mathcal{C}^\alpha(R,r,\zeta)$ and $\mathcal{D}^\beta(R,\gamma)$

$$P\big(|V_{b,X}| \ge \sqrt{\kappa_n}\big(\sqrt{2}\sigma_{b,X} + \mathrm{o}\big(n^{-1/2}\big(b^{-\beta+1/2} \vee 1\big)\big)\big)\big) \le 2\mathrm{e}^{-\kappa_n}.$$

The result follows from choosing $\kappa = (1+\delta)^2\log\log n$ and using $|\mathcal{B}| \lesssim \log n$.

(ii) Using an estimate as in (50), we obtain

$$\begin{aligned}
|V_{b,\varepsilon}| &\le \frac{1}{2\pi}\left(\int_{\mathbb{R}} |\varphi_K(bu)|\left|\frac{\varphi_n(u)}{\varphi_\varepsilon(u)}\right|^2 \mathrm{d}u\right)^{1/2}\left(\int_{\mathbb{R}} |\varphi_K(bu)|\big|\mathcal{F}a_s(-u)\big|^2\left|\frac{\varphi_\varepsilon(u) - \varphi_{\varepsilon,m}(u)}{\varphi_{\varepsilon,m}(u)}\right|^2 \mathrm{d}u\right)^{1/2} \\
&\le \frac{1}{2\pi}\underbrace{\left(\int_{\mathbb{R}} |\varphi_K(bu)|\left|\frac{\varphi_n(u)}{\varphi_\varepsilon(u)}\right|^2 \mathrm{d}u\right)^{1/2}}_{=:V_{b,\varepsilon,1}} \\
&\quad\times \underbrace{\left(\int_{\mathbb{R}} |\varphi_K(bu)|\left|\frac{\mathcal{F}a_s(-u)}{\varphi_{\varepsilon,m}(u)}\right|^2 \mathrm{d}u\right)^{1/2}}_{=:V_{b,\varepsilon,2}} \sup_{|u|\le 1/b}\big|\varphi_\varepsilon(u) - \varphi_{\varepsilon,m}(u)\big|.
\end{aligned}$$

Hence, for any $c \in (0, 1/4)$

$$\begin{aligned}
&P\big(\big\{|V_{b,\varepsilon}| \ge \delta(\log n)^3\sigma_{b,\varepsilon}\big\} \cap B_\varepsilon(b)\big) \\
&\quad\le P\big(|V_{b,\varepsilon,1}| \ge (\log n)^{1+c}\sigma_{b,\varepsilon,1}\big) + P\big(\big\{|V_{b,\varepsilon,2}| \ge (\log n)^{1+c}\sigma_{b,\varepsilon,2}\big\} \cap B_\varepsilon(b)\big) \\
&\qquad + P\Big(\sup_{|u|\le 1/b}\big|\varphi_\varepsilon(u) - \varphi_{\varepsilon,m}(u)\big| \ge \delta(\log n)^{1-2c}m^{-1/2}\Big) \\
&\quad =: P_{b,1} + P_{b,2} + P_{b,3}.
\end{aligned}$$

The first two probabilities can be bounded by Markov's inequality:

$$P_{b,1} \le (\log n)^{-1-c}\sigma_{b,\varepsilon,1}^{-1}\mathbb{E}[V_{b,\varepsilon,1}] = (\log n)^{-1-c},$$

$$P_{b,2} \le (\log n)^{-1-c}\sigma_{b,\varepsilon,2}^{-1}\mathbb{E}[V_{b,\varepsilon,2}\mathbb{1}_{B_\varepsilon(b)}] = (\log n)^{-1-c}.$$

For $P_{b,3}$ we will apply the following version of Talagrand's inequality (cf. Massart [19], (5.50)): let T be a countable index and for all $t \in T$ let $Z_{1,t}, \dots, Z_{n,t}$ be an i.i.d. sample of centered, complex valued random variables satisfying $\|Z_{k,t}\|_\infty \le b$, for all $t \in T, k = 1, \dots, n$, as well as $\sup_{t\in T} \mathrm{Var}(\sum_{k=1}^n Z_{k,t}) \le v < \infty$. Then for all $\kappa > 0$

$$P\left(\sup_{t\in T}\left|\sum_{k=1}^n Z_{k,t}\right| \ge 4\mathbb{E}\left[\sup_{t\in T}\left|\sum_{k=1}^n Z_{k,t}\right|\right] + \sqrt{2v\kappa} + \frac{2}{3}b\kappa\right) \le 2\mathrm{e}^{-\kappa}. \tag{52}$$

Choosing the rational numbers $T = \mathbb{Q}\cap[-\frac{1}{b}, \frac{1}{b}]$ and $Z_{k,t} := \mathrm{e}^{\mathrm{i}t\varepsilon_k^*} - \varphi_\varepsilon(t)$, Talagrand's inequality applies with $b = 2$ and $v = n$. As in (14), we use Theorem 4.1 by Neumann and Reiß [23] to obtain for any $\eta \in (0, 1/2)$

$$m^{1/2}\mathbb{E}\Big[\sup_{|u|\le 1/b}\big|\varphi_{\varepsilon,m}(t) - \varphi_\varepsilon(t)\big|\Big] \lesssim |\log b|^{1/2+\eta}.$$

Therefore on the assumptions $\kappa_n^{-1}(\log n)^{1+2\eta} \to 0$ and $\kappa_n/m \to 0$

$$4\mathbb{E}\Big[\sup_{|u|\le 1/b, u\in\mathbb{Q}}\big|\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)\big|\Big] + \sqrt{\frac{2\kappa_n}{m}} + \frac{4}{3m}\kappa_n = \sqrt{\frac{\kappa_n}{m}}\big(\sqrt{2} + \mathrm{o}(1)\big)$$

and thus continuity of $\varphi_{\varepsilon,m}$ and (52) yield

$$P_{b,3} = P\Big(\sup_{|u|\le 1/b, u\in\mathbb{Q}}\big|\varphi_{\varepsilon,m}(u) - \varphi_\varepsilon(u)\big| \ge \big(\sqrt{2} + \mathrm{o}(1)\big)\sqrt{\kappa_n/m}\Big) \le 2\mathrm{e}^{-\kappa_n}. \tag{53}$$

With $\kappa_n = \frac{\delta}{2}(\log n)^{2-4c}$ for $c < 1/4 - \eta/2$, we obtain $P_3 \le 2n^{-\delta/2}$. Using $b_1 = \min\mathcal{B}, |B| \lesssim \log n$ and Lemma 5.1, we finally get

$$P\Big(\sup_{b\in\mathcal{B}}|V_{b,\varepsilon}| \ge (\sqrt{2}+\delta)(\log n)^3\sigma_{b,\varepsilon}\Big) \le \sum_{b\in\mathcal{B}}(P_{b,1} + P_{b,2} + P_{b,3}) + P\big(B_\varepsilon(b_1)^c\big) = \mathrm{o}(1).$$

(iii) Corollary 5.3 shows for $\delta_b > 0$ and for any sequence $(x_n)_n$ that tends to infinity

$$P\big(\exists b \in \mathcal{B}\colon\ |V_{b,c}| \ge \delta_b\big) \lesssim \sum_{b\in\mathcal{B}} \frac{x_n}{\delta_b^2 n(mb^{2\beta+2}\wedge 1)} + \mathrm{o}(1).$$

Choosing $\delta_b = (\log n)^{3/2}n^{-1/2}(b^{-\beta+1/2} \vee 1)$ and $x_n = \mathrm{o}((\log n)^{1/2})$ yields

$$\begin{aligned} &P\big(\exists b \in \mathcal{B}\colon\ |V_{b,c}| \ge (\log n)^{3/2}n^{-1/2}\big(b^{-\beta+1/2}\vee 1\big)\big) \\ &\quad\lesssim \sum_{b\in\mathcal{B}} \frac{x_n}{(\log n)^3(mb^{(2\beta\wedge 1)+2}\wedge 1)} + \mathrm{o}(1) \lesssim \frac{x_n}{(\log n)^2(mb^{(2\beta\wedge 1)+2}\wedge 1)} + \mathrm{o}(1) = \mathrm{o}(1). \end{aligned}$$

□

For the denominator in the error representation (47) we need uniform consistency. A uniform result on the error $|\widetilde{q}_{\tau,b} - q_\tau|$ follows immediately.

Lemma 5.6. *Let $\mathcal{B}$ be a finite set satisfying $|\mathcal{B}| \lesssim \log n$, $\sup_{b\in\mathcal{B}} b\log(n) \to 0$ as well as $\sup_{b\in\mathcal{B}}(\log n)^2/(nb^{2\beta+1}) \to 0$. Then we obtain for $n \to \infty$ and $\eta \in (0,1)$*

$$\sup_{f\in\mathcal{C}^\alpha(R,r,\zeta,U_n)} \sup_{f_\varepsilon\in\mathcal{D}^\beta(R,\gamma)} P\Big(\sup_{b\in\mathcal{B}} \sup_{q_\tau^*\in[q_\tau\wedge\widetilde{q}_{\tau,b},q_\tau\vee\widetilde{q}_{\tau,b}]} \big|\widetilde{f}_b\big(q_\tau^*\big) - f(q_\tau)\big| > \eta f(q_\tau)\Big) \to 0. \tag{54}$$

Moreover, supposing $\min_{b\in\mathcal{B}} nb^{(2\beta\wedge 1)+2} \gtrsim 1$, we obtain uniformly in $f \in \mathcal{C}^\alpha(R,r,\zeta)$ and $f_\varepsilon \in \mathcal{D}^\beta(R,\gamma)$ for any sequence of critical values $(\delta_b)_{b\in\mathcal{B}}$ satisfying $\inf_{\mathcal{B}}\delta_b \to \infty$

$$P\big(\exists b\in\mathcal{B}\colon |\widetilde{q}_{\tau,b} - q_\tau| > \delta_b\big(3Db^{\alpha+1} + n^{-1/2}\big(b^{-\beta+1/2}\vee 1\big)\big)\big) \lesssim \sum_{b\in\mathcal{B}} \frac{1}{\delta_b} + \mathrm{o}(1). \tag{55}$$

Proof. Since $f(q_\tau) \ge r$ and $f \in C^\alpha([q_\tau - \zeta, q_\tau + \zeta], R)$, decomposition (35) implies with $\kappa = (\frac{\eta r}{2R})^{1\vee\alpha^{-1}} \wedge \zeta$

$$\begin{aligned} &P\Big(\sup_{b\in\mathcal{B}} \sup_{q_\tau^*\in[q_\tau\wedge\widetilde{q}_{\tau,b},q_\tau\vee\widetilde{q}_{\tau,b}]} \big|\widetilde{f}_b\big(q_\tau^*\big) - f(q_\tau)\big| > \eta f(q_\tau)\Big) \\ &\quad\le P\Big(\sup_{b\in\mathcal{B}} \sup_{x\in[-\kappa,\kappa]} \big|\widetilde{f}_b(x+q_\tau) - f(x+q_\tau)\big| > \eta r/2\Big) + P\Big(\sup_{b\in\mathcal{B}} |\widetilde{q}_{\tau,b} - q_\tau| > \kappa\Big). \end{aligned} \tag{56}$$

Using $b_1 = \min\mathcal{B}$, the first probability can be bounded by

$$\begin{aligned} &\sum_{b\in\mathcal{B}} P\Big(\Big\{\sup_{x\in[-\kappa,\kappa]} \big|\widetilde{f}_b(x+q_\tau) - f(x+q_\tau)\big| > \eta r/2\Big\} \cap B_\varepsilon(b_1)\Big) + P\big(B_\varepsilon(b_1)^c\big) \\ &\quad\lesssim \log n \sup_{b\in\mathcal{B}} P\Big(\Big\{\sup_{x\in[-\kappa,\kappa]} \big|\widetilde{f}_b(x+q_\tau) - f(x+q_\tau)\big| > \eta r/2\Big\} \cap B_\varepsilon(b_1)\Big) + \mathrm{o}(1) = \mathrm{o}(1), \end{aligned}$$

since for all b the probability in the last line converges faster to zero than $1/\log n$ owing to the concentration inequalities (27) and (28) and the conditions on b. To estimate the second term in (56), we apply Lemma 5.4. Therefore, the conditions $b\log(n) \to 0$ and $(\log n)^2/(nb^{2\beta+1}) \to 0$ yield the first assertion.

The estimate (55) follows from the error decomposition (4), (54) and Corollary 5.3 with $x_n = \mathrm{o}(\inf_{\mathcal{B}}\delta_b)$:

$$\begin{aligned} &P\big(\exists b\in\mathcal{B}\colon |\widetilde{q}_{\tau,b} - q_\tau| > \delta_b\big(3Db^{\alpha+1} + n^{-1/2}\big(b^{-\beta+1/2}\vee 1\big)\big)\big) \\ &\quad\le P\Big(\exists b\in\mathcal{B}\colon \Big|\int_{-\infty}^{q_\tau} \widetilde{f}_b(x) - f(x)\,\mathrm{d}x\Big| > \frac{1}{2} f(q_\tau)\delta_b\big(3Db^{\alpha+1} + n^{-1/2}\big(b^{-\beta+1/2}\vee 1\big)\big)\Big) \end{aligned}$$

$$
\begin{aligned}
&+ P\Big(\sup_{b\in\mathcal{B}}\sup_{q_\tau^*\in[q_\tau\wedge\widetilde{q}_{\tau,b},q_\tau\vee\widetilde{q}_{\tau,b}]}\big|\widetilde{f}_b\big(q_\tau^*\big)-f(q_\tau)\big|>\frac{1}{2}f(q_\tau)\Big)\\
&\lesssim \sum_{b\in\mathcal{B}}\Big(\frac{1}{\delta_b}+\frac{1}{\delta_b^2}\frac{x_n}{mb^{1\wedge 2\beta+2}\wedge 1}\Big)+\mathrm{o}(1)\lesssim\sum_{b\in\mathcal{B}}\frac{1}{\delta_b}+\mathrm{o}(1). \qquad \square
\end{aligned}
$$

The variances $\sigma_{b,X}$ and $\sigma_{b,\varepsilon}$, defined in (49) and (51) can be estimated by $\widetilde{\sigma}_{b,X}$ and $\widetilde{\sigma}_{b,\varepsilon}$ from (11) and (12), respectively. The latter can be decomposed into $\widetilde{\sigma}_{b,\varepsilon}^2=\frac{1}{4}\pi^{-2}m^{-1}\widetilde{\sigma}_{b,\varepsilon,1}^2\widetilde{\sigma}_{b,\varepsilon,2}^2$ with

$$
\begin{aligned}
\widetilde{\sigma}_{b,\varepsilon,1}^2&=\int_{-1/b}^{1/b}\big|\varphi_K(bu)\big|\bigg|\frac{\varphi_n(u)}{\varphi_{\varepsilon,m}(u)}\bigg|^2\,\mathrm{d}u,\\
\widetilde{\sigma}_{b,\varepsilon,2}^2&=\int_{-1/b}^{1/b}\big|\varphi_K(bu)\big|\frac{|\mathcal{F}a_s(u)|^2}{|\varphi_{\varepsilon,m}|^2}\,\mathrm{d}u.
\end{aligned}
$$

The following two lemmas show that these estimators are indeed reasonable.

Lemma 5.7. *Let $\mathcal{B}$ be a finite set satisfying $|\mathcal{B}|\lesssim\log n$, $\max_{b\in\mathcal{B}}b^\alpha\log n\to 0$ as well as $\min_{b\in\mathcal{B}}nb^{2\beta+2}\to\infty$. Let $\widetilde{\sigma}_{b,X}$ and $\sigma_{b,X}$ be given in* (11) *and* (49)*, respectively. Then we obtain for all $\eta>0$ as $n\to\infty$*

$$
\sup_{f\in\mathcal{C}^\alpha(R,r,\zeta)}\sup_{f_\varepsilon\in\mathcal{D}^\beta(R,\gamma)}P\big(\exists b\in\mathcal{B}\colon|\widetilde{\sigma}_{b,X}-\sigma_{b,X}|>\eta m^{-1/2}\big(b^{-\beta+1/2}\vee 1\big)\big)\to 0.
$$

Proof. Note that

$$
\begin{aligned}
\widetilde{\sigma}_{b,X}^2&=\frac{1}{n^2}\sum_{j=1}^n\xi_{j,1}^2(b)+\frac{1}{n^2}\sum_{j=1}^n\xi_{j,2}^2(b)+\frac{1}{n^2}\sum_{j=1}^n\xi_{j,3}^2(b)\\
&\quad+\frac{2}{n^2}\sum_{j=1}^n\xi_{j,1}(b)\xi_{j,2}(b)+\frac{2}{n^2}\sum_{j=1}^n\xi_{j,1}(b)\xi_{j,3}(b)+\frac{2}{n^2}\sum_{j=1}^n\xi_{j,2}(b)\xi_{j,3}(b),
\end{aligned}
\tag{57}
$$

where we have defined

$$
\begin{aligned}
\xi_{j,1}(b)&:=\int_{-\infty}^0 a_s(x)\mathcal{F}^{-1}\bigg[\varphi_K(bu)\mathrm{e}^{\mathrm{i}uY_j}\bigg(\frac{1}{\varphi_{\varepsilon,m}(u)}-\frac{1}{\varphi_\varepsilon(u)}\bigg)\bigg](x+\widetilde{q}_{\tau,b})\,\mathrm{d}x,\\
\xi_{j,2}(b)&:=\int_{-\infty}^0 a_s(x)\mathcal{F}^{-1}\bigg[\frac{\varphi_K(bu)\mathrm{e}^{\mathrm{i}uY_j}}{\varphi_\varepsilon(u)}\bigg](x+q_\tau)\,\mathrm{d}x,\\
\xi_{j,3}(b)&:=\int_{-\infty}^0 a_s(x)\mathcal{F}^{-1}\bigg[\frac{\varphi_K(bu)\mathrm{e}^{\mathrm{i}uY_j}(\mathrm{e}^{-\mathrm{i}u\widetilde{q}_{\tau,b}}-\mathrm{e}^{-\mathrm{i}uq_\tau})}{\varphi_\varepsilon(u)}\bigg](x)\,\mathrm{d}x.
\end{aligned}
$$

We will first study these three terms separately. Applying Plancherel's identity, the Cauchy–Schwarz inequality, the Neumann type bound (25) as well as $|\mathcal{F}a_s(u)|\le A_s(1+|u|)^{-1}$, the decay of φ_ε and the upper bound on f, we obtain

$$\begin{aligned}\mathbb{E}\bigl[\bigl|\xi_{j,1}(b)\bigr|^2\mathbb{1}_{B_\varepsilon(b)}\bigr] &\le \frac{9}{2\pi^2}\int_{-1/b}^{1/b}\frac{|\mathcal{F}a_s(u)|^2}{|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u\int_{-1/b}^{1/b}\frac{|\varphi_K(bu)|^2}{m|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u \\ &\lesssim \frac{1}{(b^{2\beta-1}\wedge 1)mb^{2\beta+1}},\end{aligned} \tag{58}$$

$$\begin{aligned}\mathbb{E}\bigl[\bigl|\xi_{j,2}(b)\bigr|^2\bigr] &= \mathbb{E}\biggl[\biggl|\frac{1}{2\pi}\int_{\mathbb{R}}\mathcal{F}a_s(u)\mathrm{e}^{-\mathrm{i}uq_\tau}\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)}\mathrm{e}^{\mathrm{i}uY_j}\,\mathrm{d}u\biggr|^2\biggr] \\ &\le \frac{\|K\|_{L^1}^2A_s^2R^3}{4\pi^2}\int_{-1/b}^{1/b}\bigl(1+|u|\bigr)^{2\beta-2}\,\mathrm{d}u =: S_b^2\end{aligned} \tag{59}$$

as well as the deterministic bound

$$\bigl|\xi_{j,2}(b)\bigr|^2 = \biggl|\frac{1}{2\pi}\int_{\mathbb{R}}\mathcal{F}a_s(u)\mathrm{e}^{-\mathrm{i}uq_\tau}\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)}\mathrm{e}^{\mathrm{i}uY_j}\,\mathrm{d}u\biggr|^2 \le \frac{\|K\|_{L_1}^2A_s^2}{4\pi^2}\int_{-1/b}^{1/b}\bigl(1+|u|\bigr)^{2\beta}\,\mathrm{d}u =: d_b^2.$$

Hence, $\mathrm{Var}[\xi_{j,2}(b)^2] \le \mathbb{E}[\xi_{j,2}(b)^4] \le d_b^2S_b^2$ and $|\xi_{j,2}^2(b) - \mathbb{E}[\xi_{j,2}^2(b)]| \le 2d_b^2$, so that an application of Bernstein's inequality yields for any $b > 0$ and $z > 0$

$$P\Biggl(\Biggl|\frac{1}{n}\sum_{j=1}^n\bigl(\xi_{j,2}^2(b) - \mathbb{E}\bigl[\xi_{j,2}^2(b)\bigr]\bigr)\Biggr| \ge z\Biggr) \le 2\exp\biggl(-\frac{z^2n}{2S_b^2d_b^2 + (4/3)d_b^2z}\biggr).$$

Setting $z = S_b^2$ and noting $S_b^2 \lesssim (b^{-2\beta+1}\vee 1)$, $d_b^2 \lesssim b^{-2\beta}$, we see that

$$P\Biggl(\Biggl|\frac{1}{n}\sum_{j=1}^n\bigl(\xi_{j,2}^2(b) - \mathbb{E}\bigl[\xi_{j,2}^2(b)\bigr]\bigr)\Biggr| \ge S_b^2\Biggr) \le 2\exp\biggl(-\frac{S_b^2n}{4d_b^2}\biggr) \le 2\exp\bigl(-Cnb^{2\beta\wedge 1}\bigr) \tag{60}$$

for some $C > 0$. The right-hand side of (60) tends to zero with polynomial rate since $nb^{2\beta\wedge 1} \gtrsim \log n$.

We use $\operatorname{supp} a_s \subseteq [-1, 0]$ to write $\xi_{j,3}$ as

$$\begin{aligned}\xi_{j,3}(b) &= \int_{\mathbb{R}}\bigl(a_s(x-\widetilde{q}_{\tau,b}) - a_s(x-q_\tau)\bigr)\mathcal{F}^{-1}\biggl[\frac{\varphi_K(bu)\mathrm{e}^{\mathrm{i}uY_j}}{\varphi_\varepsilon(u)}\biggr](x)\,\mathrm{d}x \\ &\le \sup_{t\in(-1,0)}\bigl|a_s'(t)\bigr||\widetilde{q}_{\tau,b} - q_\tau|\int_{(\widetilde{q}_{\tau,b}\wedge q_\tau)-1}^{\widetilde{q}_{\tau,b}\vee q_\tau}\biggl|\mathcal{F}^{-1}\biggl[\frac{\varphi_K(bu)\mathrm{e}^{\mathrm{i}uY_j}}{\varphi_\varepsilon(u)}\biggr](x)\biggr|\,\mathrm{d}x.\end{aligned}$$

The Cauchy–Schwarz inequality and Plancherel's identity yield

$$\begin{aligned}\bigl|\xi_{j,3}(b)\bigr|^2 \le{}& \bigl\|a_s'\mathbb{1}_{(-1,0)}\bigr\|_\infty^2|\widetilde{q}_{\tau,b} - q_\tau|^2\bigl(1 + |\widetilde{q}_{\tau,b} - q_\tau|\bigr) \\ &\times\int_{(\widetilde{q}_{\tau,b}\wedge q_\tau)-1}^{\widetilde{q}_{\tau,b}\vee q_\tau}\biggl|\mathcal{F}^{-1}\biggl[\frac{\varphi_K(bu)\mathrm{e}^{\mathrm{i}uY_j}}{\varphi_\varepsilon(u)}\biggr](x)\biggr|^2\,\mathrm{d}x\end{aligned}$$

$$\leq \frac{\|a_s'\mathbb{1}_{(-1,0)}\|_\infty^2}{2\pi}|\widetilde{q}_{\tau,b}-q_\tau|^2\big(1+|\widetilde{q}_{\tau,b}-q_\tau|\big)\int_{\mathbb{R}}\left|\frac{\varphi_K(bu)}{\varphi_\varepsilon(u)}\right|^2\mathrm{d}u$$
$$\lesssim |\widetilde{q}_{\tau,b}-q_\tau|^2\big(1+|\widetilde{q}_{\tau,b}-q_\tau|\big)b^{-2\beta-1}.$$

By Lemma 5.4 $\sup_{b\in\mathcal{B}}|\widetilde{q}_{\tau,b}-q_\tau|=\mathrm{o}_P(1)$. Applying (55), we conclude for some constant $C>0$, for $\delta_b=(b^{\alpha+(1/2-\beta)_+}+n^{-1/2}b^{-\beta-1/2}))^{-1}$ and for any $\eta>0$

$$\begin{aligned}
&P\big(\exists b\in\mathcal{B}\colon\ \big|\xi_{j,3}(b)\big|>\eta\big(b^{-\beta+1/2}\vee 1\big)\big)\\
&\quad\leq P\big(\exists b\in\mathcal{B}\colon\ |\widetilde{q}_{\tau,b}-q_\tau|>\eta C b^{(\beta\wedge 1/2)+1/2}\big)+\mathrm{o}(1)\\
&\quad\leq P\big(\exists b\in\mathcal{B}\colon\ |\widetilde{q}_{\tau,b}-q_\tau|>\eta C\delta_b\big(b^{\alpha+1}+n^{-1/2}\big(b^{-\beta+1/2}\vee 1\big)\big)\big)+\mathrm{o}(1)\\
&\quad\lesssim\left(\sum_{b\in\mathcal{B}}(\delta_b)^{-1}\right)+\mathrm{o}(1)\lesssim\sup_{b\in\mathcal{B}}b^\alpha\log n+\sup_{b\in\mathcal{B}}\frac{\log n}{\sqrt{n}b^{\beta+1/2}}+\mathrm{o}(1)=\mathrm{o}(1).
\end{aligned}\tag{61}$$

Combining the variance bounds (58), (59) and (61), we apply Markov's inequality, the Cauchy–Schwarz inequality and the concentration result (60) on the decomposition (57) to obtain

$$\begin{aligned}
&\sup_{b\in\mathcal{B}}\big(n\big(b^{2\beta-1}\wedge 1\big)\big|\widetilde{\sigma}_{b,X}^2-\sigma_{b,X}^2\big|\big)\\
&\quad=\sup_{b\in\mathcal{B}}\left(\frac{b^{2\beta-1}\wedge 1}{n}\sum_{j=1}^n\big(\xi_{j,2}^2(b)-\mathbb{E}\big[\xi_{j,2}^2(b)\big]\big)\right)+\mathrm{o}_P(1)=\mathrm{o}_P(1).
\end{aligned}$$
□

Lemma 5.8. *Let $\mathcal{B}$ be a finite set satisfying $|\mathcal{B}|\lesssim\log n$ as well as $\sup_{b\in\mathcal{B}}1/(nb^{2\beta+1})\to 0$. Let $\widetilde{\sigma}_{b,\varepsilon}$ and $\sigma_{b,\varepsilon}$ be given in (12) and (51), respectively. Then we obtain uniformly over $f\in\mathcal{C}^\alpha(R,r,\zeta)$ and $f_\varepsilon\in\mathcal{D}^\beta(R,\gamma)$ for all $\eta>0$ as $n\to\infty$*

$$P\big(\exists b\in\mathcal{B}\colon\ |\widetilde{\sigma}_{b,\varepsilon}-\sigma_{b,\varepsilon}|>\eta(\log n)m^{-1/2}\big(b^{-\beta+1/2}\vee 1\big)\big)\to 0.$$

Proof. We start by showing for $b_1=\min\mathcal{B}$ that

$$\sup_{|u|\leq 1/b_1}\left|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)}\right|=1+\mathrm{o}_P(1).\tag{62}$$

To this end, recall $w(u)=(\log(\mathrm{e}+|u|))^{-1/2-\eta}$ for some $\eta\in(0,1/2)$. Markov's inequality, Lemma 5.1 and Theorem 4.1 by Neumann and Reiß [23] yield for any $\delta>0$

$$\begin{aligned}
&P\left(\sup_{|u|\leq 1/b_1}\left|\frac{\varphi_\varepsilon(u)}{\varphi_{\varepsilon,m}(u)}-1\right|\geq\delta\right)\\
&\quad\leq P\Big(\sup_{|u|\leq 1/b_1}m^{1/2}\big|\varphi_\varepsilon(u)-\varphi_{\varepsilon,m}(u)\big|\geq\delta|\log b_1|\Big)+P\Big(\inf_{|u|\leq 1/b_1}\big|\varphi_{\varepsilon,m}(u)\big|\leq m^{-1/2}|\log b_1|\Big)\\
&\quad\leq\big(\delta|\log b_1|\big)^{-1}\mathbb{E}\Big[\sup_{|u|\leq 1/b_1}m^{1/2}\big|\varphi_\varepsilon(u)-\varphi_{\varepsilon,m}(u)\big|\Big]+\mathrm{o}(1)
\end{aligned}$$

$$\leq \frac{1}{\delta|\log b_1|w(1/b_1)}\mathbb{E}\Big[\sup_{u\in\mathbb{R}} m^{1/2}w(u)\big|\varphi_\varepsilon(u)-\varphi_{\varepsilon,m}(u)\big|\Big]+\mathrm{o}(1)=\mathrm{o}(1),$$

which implies (62) holding uniformly in $\mathcal{B}$ since $[-1/b_1, 1/b_1]$ is the maximal interval for all $b\in\mathcal{B}$.

Now, we consider $\widetilde{\sigma}_{b,\varepsilon,1}$. The uniform consistency (62) implies

$$\widetilde{\sigma}^2_{b,\varepsilon,1}=\big(1+\mathrm{o}_P(1)\big)\int_{\mathbb{R}}\big|\varphi_K(bu)\big|\left|\frac{\varphi_n(u)}{\varphi_\varepsilon(u)}\right|^2\mathrm{d}u.$$

Chebyshev's inequality yields for all $\eta>0$

$$\begin{aligned}
&P\left(\sup_{b\in\mathcal{B}}\left|\left(\int_{\mathbb{R}}\big|\varphi_K(bu)\big|\frac{|\varphi_n(u)|^2}{|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u\right)^{1/2}-\mathbb{E}\left[\left(\int_{\mathbb{R}}\big|\varphi_K(bu)\big|\frac{|\varphi_n(u)|^2}{|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u\right)^{1/2}\right]\right|>\eta\log n\right)\\
&\quad\leq(\eta\log n)^{-2}\sum_{b\in\mathcal{B}}\mathbb{E}\left[\int_{\mathbb{R}}\big|\varphi_K(bu)\big|\frac{|\varphi_n(u)|^2}{|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u\right]\\
&\quad\lesssim\big(\eta^2\log n\big)^{-1}\int_{-1/b_1}^{1/b_1}\frac{\mathbb{E}[|\varphi_n(u)|^2]}{|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u\lesssim\big(\eta^2\log n\big)^{-1},
\end{aligned}$$

where the last estimate follows from $\mathbb{E}[|\varphi_n(u)|^2]\lesssim|\varphi_Y(u)|^2+\mathbb{E}[|\varphi_n(u)-\varphi_Y(u)|^2]\lesssim|\varphi_Y(u)|^2+1/n$, $f_\varepsilon\in\mathcal{D}^\beta(R,\gamma)$, $\|f\|_\infty\lesssim 1$ and $nb_1^{2\beta+1}\to\infty$. Hence, we obtain uniformly in $\mathcal{B}$

$$\widetilde{\sigma}_{b,\varepsilon,1}=\big(1+\mathrm{o}_P(1)\big)\big(\sigma_{b,\varepsilon,1}+\mathrm{o}_P(\log n)\big)=\sigma_{b,\varepsilon,1}+\mathrm{o}_P(\log n). \tag{63}$$

Concerning $\widetilde{\sigma}_{b,\varepsilon,2}$, we write with use of (62)

$$\widetilde{\sigma}^2_{b,\varepsilon,2}=\int_{-1/b}^{1/b}\big|\varphi_K(bu)\big|\frac{|\mathcal{F}a_s(u)|^2}{|\varphi_{\varepsilon,m}(u)|^2}\,\mathrm{d}u=\big(1+\mathrm{o}_p(1)\big)\int_{-1/b}^{1/b}\big|\varphi_K(bu)\big|\frac{|\mathcal{F}a_s(u)|^2}{|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u.$$

Moreover, the triangle inequality for the L^2-norm and Lemma 5.1, applied on $B_\varepsilon(b_1)$ yield

$$\begin{aligned}
&\left|\left(\int_{-1/b}^{1/b}\big|\varphi_K(bu)\big|\frac{|\mathcal{F}a_s(u)|^2}{|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u\right)^{1/2}-\sigma_{b,\varepsilon,2}\right|^2\\
&\quad\leq 2\left|\mathbb{E}\left[\left(\left(\int_{-1/b}^{1/b}\big|\varphi_K(bu)\big|\frac{|\mathcal{F}a_s(u)|^2}{|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u\right)^{1/2}\right.\right.\right.\\
&\qquad\left.\left.\left.-\left(\int_{-1/b}^{1/b}\big|\varphi_K(bu)\big|\frac{|\mathcal{F}a_s(u)|^2}{|\varphi_{\varepsilon,m}(u)|^2}\,\mathrm{d}u\right)^{1/2}\right)\mathbb{1}_{B_\varepsilon(b_1)}\right]\right|^2\\
&\qquad+2P\big(\big(B_\varepsilon(b_1)\big)^c\big)\int_{-1/b}^{1/b}\big|\varphi_K(bu)\big|\frac{|\mathcal{F}a_s(u)|^2}{|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u\\
&\quad\leq 2\mathbb{E}\left[\left(\int_{-1/b}^{1/b}\big|\varphi_K(bu)\big|\big|\mathcal{F}a_s(u)\big|^2\frac{|\varphi_{\varepsilon,m}(u)-\varphi_\varepsilon(u)|^2}{|\varphi_\varepsilon(u)\varphi_{\varepsilon,m}(u)|^2}\,\mathrm{d}u\right)\mathbb{1}_{B_\varepsilon(b_1)}\right]+\mathrm{o}(1)\int_{-1/b}^{1/b}\frac{|\mathcal{F}a_s(u)|^2}{|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u
\end{aligned}$$

$$\leq \frac{2}{|\log b_1|^{3/2}}\mathbb{E}\left[\int_{-1/b}^{1/b} \frac{|\mathcal{F}a_s(u)|^2}{|\varphi_\varepsilon(u)|^2} m\big|\varphi_{\varepsilon,m}(u)-\varphi_\varepsilon(u)\big|^2\,\mathrm{d}u\right] + \mathrm{o}(1)\big(b^{-2\beta+1}\vee 1\big)$$
$$= \mathrm{o}(1)\big(b^{-2\beta+1}\vee 1\big),$$

where $\mathrm{o}(1)$ is a null sequence which does not depend on b. Consequently,

$$\sup_{b\in\mathcal{B}}\left|\left(\int_{-1/b}^{1/b}\big|\varphi_K(bu)\big|\frac{|\mathcal{F}a_s(u)|^2}{|\varphi_\varepsilon(u)|^2}\,\mathrm{d}u\right)^{1/2} - \sigma_{b,\varepsilon,2}\right|\big(b^{\beta-1/2}\wedge 1\big) = \mathrm{o}(1).$$

Using $\sigma^2_{b,\varepsilon,2} \lesssim b^{-2\beta+1}\vee 1$ by the analysis of the convergence rates, we get

$$\widetilde{\sigma}_{b,\varepsilon,2} = \big(1+\mathrm{o}_p(1)\big)\big(\sigma_{b,\varepsilon,2} + \mathrm{o}\big(b^{-\beta+1/2}\vee 1\big)\big) = \sigma_{b,\varepsilon,2} + \mathrm{o}_P\big(b^{-\beta+1/2}\vee 1\big). \tag{64}$$

Since $\sigma_{b,\varepsilon,1}\lesssim 1$, $\sigma_{b,\varepsilon,2}\lesssim b^{-\beta+1/2}\vee 1$, it remains to combine (63) and (64) to obtain uniformly in $\mathcal{B}$

$$\widetilde{\sigma}_{b,\varepsilon} = \frac{1}{2\pi}m^{-1/2}\widetilde{\sigma}_{b,\varepsilon,1}\widetilde{\sigma}_{b,\varepsilon,2} = \frac{1}{2\pi}m^{-1/2}\big(\sigma_{b,\varepsilon,1}+\mathrm{o}_P(\log n)\big)\big(\sigma_{b,\varepsilon,2}+\mathrm{o}_P\big(b^{-\beta+1/2}\vee 1\big)\big)$$
$$= \sigma_{b,\varepsilon} + \mathrm{o}_P\big((\log n)m^{-1/2}\big(b^{-\beta+1/2}\vee 1\big)\big). \qquad \square$$

5.2.3. *Proof of Theorem* 3.2

Applying Lemma 5.1 and (44), it suffices to consider the event

$$A_0 := \big\{b_{j_{0,n}} \leq b_{\widetilde{j}_n} \leq n^{-1/(2\alpha+2(\beta\vee 1/2)+1)}\big\}\cap B_\varepsilon(b_{j_{0,n}})$$

with $j_{0,n}$ defined in (43). Therefore we can set $\mathcal{B} := \{b_{j_{0,n}},\dots,b_{M_n}\}$ in the following.

As seen in error decomposition (47), there are three stochastic errors $V_{b,X}$, $V_{b,\varepsilon}$ and $V_{b,c}$ which were treated in Lemma 5.5. This motivates the following definition. For $\delta_1>0$, let

$$S_{b,X} := (1+\delta_1)\sqrt{2\log\log n}\max_{\mu\in\mathcal{B}:\,\mu\geq b}\sigma_{\mu,X}, \qquad S_{b,\varepsilon} := (\delta_1\log n)^3\max_{\mu\in\mathcal{B}:\,\mu\geq b}\sigma_{\mu,\varepsilon}.$$

On the assumption $|\varphi_\varepsilon(u)|\gtrsim(1+|u|)^{-\beta}$ we obtain for $\sigma_{b,\varepsilon}=\frac{1}{2\pi}m^{-1/2}\sigma_{b,\varepsilon,1}\sigma_{b,\varepsilon,2}$ from (51) that

$$\sigma^2_{b,\varepsilon,2} \gtrsim \int_{-1/b}^{1/b}\big|\mathcal{F}a_s(-u)\big|^2\big(1+|u|\big)^{2\beta}\,\mathrm{d}u \gtrsim \int_{-1/b}^{1/b}\big(1+|u|\big)^{2\beta-2}\,\mathrm{d}u \sim b^{-2\beta+1}\vee 1.$$

Also, we have $\sigma_{b,\varepsilon,1} = \|\varphi_X\|_{L^2}+\mathrm{o}(1)\geq\|\varphi_X\|_{L^2}/2$ for b small enough and n large enough. Thus, $\sigma_{b,\varepsilon}\gtrsim m^{-1/2}(b^{-\beta+1/2}\vee 1)$. Therefore, Lemma 5.5 yields

$$P\big(\exists b\in\mathcal{B}\colon |V_{b,X}+V_{b,\varepsilon}+V_{b,c}|\geq S_{b,X}+S_{b,\varepsilon}\big)$$
$$\leq P\left(\exists b\in\mathcal{B}\colon |V_{b,X}|\geq S_{b,X}+\frac{1}{3}S_{b,\varepsilon}\right) + P\left(\exists b\in\mathcal{B}\colon |V_{b,\varepsilon}|\geq\frac{S_{b,\varepsilon}}{3}\right)$$

$$+ P\left(\exists b \in \mathcal{B}\colon\ |V_{b,c}| \geq \frac{S_{b,\varepsilon}}{3}\right)$$
$$= \mathrm{o}(1).$$

Hence, the probability of the event

$$A_1 := \left\{\forall b \in \mathcal{B}\colon\ |V_{b,X} + V_{b,\varepsilon} + V_{b,c}| \leq S_{b,X} + S_{b,\varepsilon}\right\}$$

converges to one. The variances $S_{b,X}$ and $S_{b,\varepsilon}$ can be estimated by

$$\widetilde{S}_{b,X} := (1+\delta_1)\sqrt{2\log\log n} \max_{\mu\in\mathcal{B}\colon \mu\geq b} \widetilde{\sigma}_{\mu,X}, \qquad \widetilde{S}_{b,\varepsilon} := (\delta_1 \log n)^3 \max_{\mu\in\mathcal{B}\colon \mu\geq b} \widetilde{\sigma}_{\mu,\varepsilon}.$$

Applying Lemmas 5.7 and 5.8, the triangle inequality of the ℓ^∞-norm yields uniformly in $b \in \mathcal{B}$

$$\left|\max_{\mu\geq b} \widetilde{\sigma}_{\mu,X} - \max_{\mu\geq b} \sigma_{\mu,X}\right| \leq \max_{\mu\geq b} |\widetilde{\sigma}_{\mu,X} - \sigma_{\mu,X}| = \mathrm{o}_P\left(\frac{1}{m^{1/2}(b^{\beta-1/2}\wedge 1)}\right),$$
$$\left|\max_{\mu\geq b} \widetilde{\sigma}_{\mu,\varepsilon} - \max_{\mu\geq b} \sigma_{\mu,\varepsilon}\right| \leq \max_{\mu\geq b} |\widetilde{\sigma}_{\mu,\varepsilon} - \sigma_{\mu,\varepsilon}| = \mathrm{o}_P\left(\frac{\log n}{m^{1/2}(b^{\beta-1/2}\wedge 1)}\right).$$

Using again $\sigma_{b,\varepsilon} \gtrsim m^{-1/2}(b^{-\beta+1/2}\vee 1)$, we thus obtain for all $\eta > 0$ that the event

$$A_2 := \left\{\forall b \in \mathcal{B}\colon\ \left|(\widetilde{S}_{b,X} + \widetilde{S}_{b,\varepsilon}) - (S_{b,X} + S_{b,\varepsilon})\right| \leq \eta(S_{b,X} + S_{b,\varepsilon})\right\}$$

fulfills $P(A_2) \to 1$. The same holds true for the events

$$A_3 := \left\{\forall b \in \mathcal{B}\colon \sup_{q^*\in[(q_\tau\wedge\widetilde{q}_{\tau,b})\vee(q_\tau\wedge\widetilde{q}_{\tau,b})]} \left|\widetilde{f}_b(q^*) - f(q_\tau)\right| \leq \eta f(q_\tau)\right\},$$
$$A_4 := \left\{\forall b \in \mathcal{B}\colon \sup_{q^*\in[(q_\tau\wedge\widetilde{q}_{\tau,b})\vee(q_\tau\wedge\widetilde{q}_{\tau,b})]} \left|\widetilde{f}_b(q^*) - \widetilde{f}_b(\widetilde{q}_{\tau,b})\right| \leq \eta\left|\widetilde{f}_b(\widetilde{q}_{\tau,b})\right|\right\}$$

by (54). Therefore, it is sufficient to work in the following on the event

$$A := A_0 \cap A_1 \cap A_2 \cap A_3 \cap A_4.$$

We show that the adaptive estimator $\widetilde{q}_\tau$ mimics the oracle estimator defined as follows. Recalling the estimate of the bias $B_b = Db^{\alpha+1}$, let the oracle bandwidth be defined by

$$b_* := \max\{b \in \mathcal{B}\colon\ B_b \leq S_{b,X} + S_{b,\varepsilon}\}. \tag{65}$$

Note that b_* is well-defined and unique since B_b is monoton increasing in b while $(S_{b,X} + S_{b,\varepsilon})$ is monton decreasing. We get the oracle estimator $\widetilde{q}_{\tau,b_*}$.

Since on A_4 for all $b \in \mathcal{B}$ and $q^* \in [(q_\tau \wedge \widetilde{q}_{\tau,b}) \vee (q_\tau \wedge \widetilde{q}_{\tau,b})]$

$$\left|\widetilde{f}_b(q^*)\right| \geq \left|\widetilde{f}_b(\widetilde{q}_{\tau,b})\right| - \left|\widetilde{f}_b(q^*) - \widetilde{f}_b(\widetilde{q}_{\tau,b})\right| \geq (1-\eta)\left|\widetilde{f}_b(\widetilde{q}_{\tau,b})\right|,$$

we have for any $b \in \mathcal{B}$ on the event $A_1 \cap A_4$ by (47)

$$|\widetilde{q}_{\tau,b} - q_\tau| \leq \frac{B_b + |V_{b,X} + V_{b,\varepsilon} + V_{b,c}| + |\widetilde{M}_b(\widetilde{q}_{\tau,b})|}{|\widetilde{f}_b(q^*)|} \leq \frac{B_b + S_{b,X} + S_{b,\varepsilon} + |\widetilde{M}_b(\widetilde{q}_{\tau,b})|}{(1-\eta)|\widetilde{f}_b(\widetilde{q}_{\tau,b})|}.$$

Furthermore, by the definition of b_* we have on the event A for any $b \leq b_*$

$$|\widetilde{q}_{\tau,b} - q_\tau| \leq \frac{2(S_{b,X} + S_{b,\varepsilon}) + |\widetilde{M}_b(\widetilde{q}_{\tau,b})|}{(1-\eta)|\widetilde{f}_b(\widetilde{q}_{\tau,b})|}.$$

On A_2 we estimate $\widetilde{S}_{b,X} + \widetilde{S}_{b,\varepsilon} \geq (1-\eta)(S_{b,X} + S_{b,\varepsilon})$ and thus we have on A for any $b \leq b_*$

$$|\widetilde{q}_{\tau,b} - q_\tau| \leq \frac{2(\widetilde{S}_{b,X} + \widetilde{S}_{b,\varepsilon})}{(1-\eta)^2|\widetilde{f}_b(\widetilde{q}_{\tau,b})|} + \frac{|\widetilde{M}_b(\widetilde{q}_{\tau,b})|}{(1-\eta)|\widetilde{f}_b(\widetilde{q}_{\tau,b})|}.$$

Since for any $\delta > 0$ we find $\delta_1, \eta > 0$ such that $((1-\eta)^{-2}(2\sqrt{2}+\delta_1) - 2\sqrt{2}) \vee (2(1-\eta)^{-2}\delta_1) \vee \frac{\eta}{1-\eta} < \delta$, we obtain $|\widetilde{q}_{\tau,b} - q_\tau| \leq \widetilde{\Sigma}_b$ with $\widetilde{\Sigma}_b$ as defined in (10). As a result one has $q_\tau \in \mathcal{U}_b$ and $q_\tau \in \mathcal{U}_\mu$ for all $b \leq b_*$ and $\mu \leq b_*$, implying $\mathcal{U}_\mu \cap \mathcal{U}_b \neq \varnothing$. By the definition of the procedure, $\widetilde{b}^* \geq b_*$ and $\mathcal{U}_{\widetilde{b}_*} \cap \mathcal{U}_{b_*} \neq \varnothing$ on the event A. This leads to

$$|\widetilde{q}_{\tau,\widetilde{b}^*} - q_\tau| \leq |\widetilde{q}_{\tau,b_*} - q_\tau| + |\widetilde{q}_{\tau,\widetilde{b}^*} - \widetilde{q}_{\tau,b_*}| \leq \widetilde{\Sigma}_{b_*} + (\widetilde{\Sigma}_{b_*} + \widetilde{\Sigma}_{\widetilde{b}^*}).$$

On $A_2 \cap A_3$ we have $\widetilde{\Sigma}_b \lesssim S_{b,X} + S_{b,\varepsilon}$ since $f(q_\tau) \geq r$ and $|\widetilde{M}_b(\widetilde{q}_{\tau,b})| \leq |\widetilde{M}_b(q_\tau)| = |\int_{-\infty}^{q_\tau}(\widetilde{f}_b - f)|$. Using additionally the monotonicity of $(S_{b,X} + S_{b,\varepsilon})$ as well as $\widetilde{b}^* \geq b_*$, this implies

$$|\widetilde{q}_{\tau,\widetilde{b}^*} - q_\tau| \lesssim (S_{b_*,X} + S_{b_*,\varepsilon}) \lesssim \left(\sqrt{\log\log n} + \left(\log n^\delta\right)^3\right)\left(b_*^{-\beta+1/2} \vee 1\right)n^{-1/2}.$$

It remains to note by the definition (65) of the oracle b_* and by the assumption $b_{j+1}/b_j \lesssim 1$ that $b_* \sim ((\log n^\delta)^6/n)^{-1/(2\alpha+2(\beta\vee 1/2)+1)}$ as $n \to \infty$.

Acknowledgements

This research started when the first author was a Postdoc at EURANDOM, Eindhoven University of Technology, The Netherlands. The research was partly supported by the Deutsche Forschungsgemeinschaft through the FOR 1735 “Structural Inference in Statistics”. Part of the work on this project was done during a visit of the third author to EURANDOM. We thank three anonymous referees for helpful comments and suggestions.

References

[1] Carroll, R.J. and Hall, P. (1988). Optimal rates of convergence for deconvolving a density. *J. Amer. Statist. Assoc.* **83** 1184–1186. MR0997599

[2] Carroll, R.J., Ruppert, D., Stefanski, L.A. and Crainiceanu, C.M. (2006). *Measurement Error in Nonlinear Models*: *A Modern Perspective*, 2nd ed. *Monographs on Statistics and Applied Probability* **105**. Boca Raton, FL: Chapman & Hall/CRC. MR2243417

[3] Comte, F. and Lacour, C. (2011). Data-driven density estimation in the presence of additive noise with unknown distribution. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **73** 601–627. MR2853732

[4] Dattner, I., Goldenshluger, A. and Juditsky, A. (2011). On deconvolution of distribution functions. *Ann. Statist.* **39** 2477–2501. MR2906875

[5] Dattner, I. and Reiser, B. (2013). Estimation of distribution functions in measurement error models. *J. Statist. Plann. Inference* **143** 479–493. MR2995109

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

[7] Dudley, R.M. (1992). Fréchet differentiability, p-variation and uniform Donsker classes. *Ann. Probab.* **20** 1968–1982. MR1188050

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

[9] Frese, E.M., Fick, A. and Sadowsky, H.S. (2011). Blood pressure measurement guidelines for physical therapists. *Cardiopulm. Phys. Ther. J.* **22** 5–12.

[10] Girardi, M. and Weis, L. (2003). Operator-valued Fourier multiplier theorems on Besov spaces. *Math. Nachr.* **251** 34–51. MR1960803

[11] Goldenshluger, A. and Nemirovski, A. (1997). On spatially adaptive estimation of nonparametric regression. *Math. Methods Statist.* **6** 135–170. MR1466625

[12] Hall, P. and Lahiri, S.N. (2008). Estimation of distributions, moments and quantiles in deconvolution problems. *Ann. Statist.* **36** 2110–2134. MR2458181

[13] Johannes, J. (2009). Deconvolution with unknown error distribution. *Ann. Statist.* **37** 2301–2323. MR2543693

[14] Johannes, J. and Schwarz, M. (2013). Adaptive circular deconvolution by model selection under unknown error distribution. *Bernoulli* **19** 1576–1611. MR3129026

[15] Kannel, W.B. (1995). Framingham study insights into hypertensive risk of cardiovascular disease. *Hypertension Research*: *Official Journal of the Japanese Society of Hypertension* **18** 181–196.

[16] Kappus, J. (2014). Adaptive nonparametric estimation for Lévy processes observed at low frequency. *Stochastic Process. Appl.* **124** 730–758. MR3131312

[17] Lepskiĭ, O.V. (1990). A problem of adaptive estimation in Gaussian white noise. *Teor. Veroyatn. Primen.* **35** 459–470. MR1091202

[18] Lounici, K. and Nickl, R. (2011). Global uniform risk bounds for wavelet deconvolution estimators. *Ann. Statist.* **39** 201–231. MR2797844

[19] Massart, P. (2007). *Concentration Inequalities and Model Selection*. *Lecture Notes in Math.* **1896**. Berlin: Springer. MR2319879

[20] Meister, A. (2004). On the effect of misspecifying the error density in a deconvolution problem. *Canad. J. Statist.* **32** 439–449. MR2125855

[21] Neumann, M.H. (1997). On the effect of estimating the error density in nonparametric deconvolution. *J. Nonparametr. Stat.* **7** 307–330. MR1460203

[22] Neumann, M.H. (2007). Deconvolution from panel data with unknown error distribution. *J. Multivariate Anal.* **98** 1955–1968. MR2396948

[23] Neumann, M.H. and Reiß, M. (2009). Nonparametric estimation for Lévy processes from low-frequency observations. *Bernoulli* **15** 223–248. MR2546805

[24] Nickl, R. and Reiß, M. (2012). A Donsker theorem for Lévy measures. *J. Funct. Anal.* **263** 3306–3332. MR2973342

[25] Rosenthal, H.P. (1970). On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables. *Israel J. Math*. **8** 273–303. MR0271721

[26] Söhl, J. and Trabs, M. (2012). A uniform central limit theorem and efficiency for deconvolution estimators. *Electron. J. Stat*. **6** 2486–2518. MR3020273

[27] Spokoiny, V.G. (1996). Adaptive hypothesis testing using wavelets. *Ann. Statist*. **24** 2477–2498. MR1425962

[28] Stirnemann, J.J., Comte, F. and Samson, A. (2012). Density estimation of a biomedical variable subject to measurement error using an auxiliary set of replicate observations. *Stat. Med*. **31** 4154–4163. MR3040071

[29] Triebel, H. (2010). *Theory of Function Spaces*. *Modern Birkhäuser Classics*. Basel: Birkhäuser. Reprint of 1983 edition [MR0730762], Also published in 1983 by Birkhäuser Verlag [MR0781540]. MR3024598

[30] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. New York: Springer. Revised and extended from the 2004 French original, Translated by Vladimir Zaiats. MR2724359

[31] van der Vaart, A.W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247

Received May 2013 and revised April 2014