Smoothing parameters for deconvolution recursive kernel density estimators defined by stochastic approximation method with Laplace errors

Yousri Slaoui

Université de Poitiers, France

Abstract

In this paper, we propose a bandwidth selection in deconvolution recursive kernel estimators of a probability density function defined by the stochastic approximation algorithm for Laplace errors. We show that, using the proposed bandwidth selection and the stepsize which minimize the $MISE$ (Mean Integrated Squared Error), the recursive estimator will be better than the nonrecursive one for small sample setting and when the error variance is controlled by the noise to signal ratio. We corroborate these theoretical results through simulations and a real dataset.

Key Words: Bandwidth selection, deconvolution, density estimation, plug-in methods, stochastic approximation algorithm

1 Introduction

Suppose we observe contaminated data $Y_1, \ldots, Y_n$ which are independent, identically distributed random variables, and let $f_Y$ denote the probability density of $Y_1$, where

$$Y_i = X_i + \varepsilon_i, \quad i = 1, \ldots, n$$
and where $X_1, \ldots, X_n$ are independent, identically distributed random variables, and $f_X$ denote the probability density of $X_1$. We assume that $X$ and $\varepsilon$ are mutually independent. The density function of $\varepsilon$ is denoted by $f_\varepsilon$, assumed known. Throughout this paper we suppose that $\varepsilon$ is a centred double exponentially distributed, also called Laplace distribution, and denoted by $\varepsilon \sim \mathcal{E}d(\sigma)$, with $\sigma$ is the scale parameter. Following Robbins-Monro’s procedure, we construct a stochastic algorithm, which approximates the function $f_X$ at a given point $x$, by defining an algorithm of search of the zero of the function $h : y \to f_X(x) - y$. This algorithm can be defined by setting $f_{0,X}(x) \in \mathbb{R}$, and, for all $n \geq 1$,

$$f_{n,X}(x) = f_{n-1,X}(x) + \gamma_n W_n,$$

where $W_n(x)$ is an observation of the function $h$ at the point $f_{n-1,X}(x)$, and the stepsize ($\gamma_n$) is a sequence of positive real numbers that goes to zero. To define $W_n(x)$, we follow the approach of Révész (1973, 1977), Tsybakov (1990), Mokkadem et al. (2009a, b) and of Slaoui (2014b, 2015a), and we introduce a bandwidth ($h_n$) (that is, a sequence of positive real numbers that goes to zero), and a kernel $K$ (that is, a function satisfying $\int_{\mathbb{R}} K(x) \, dx = 1$), and a deconvoluting kernel $K^\varepsilon$ defined as follows:

$$K^\varepsilon(u) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itu} \frac{\phi K(t)}{\phi_\varepsilon \left( \frac{t}{h_n} \right)} \, dt,$$

with $\phi_L$ the Fourier transform of a function or a random variable $L$, and sets $W_n(x) = h_n^{-1}K^\varepsilon(h_n^{-1}(x - Y_n)) - f_{n-1,X}(x)$. Then, the estimator $f_{n,X}$ to recursively estimate the density function $f_X$ at the point $x$ can be written as

$$f_{n,X}(x) = (1 - \gamma_n) f_{n-1,X}(x) + \gamma_n h_n^{-1}K^\varepsilon(h_n^{-1}(x - Y_n)). \quad (1)$$

This estimator was introduced by Mokkadem et al. (2009a) in the error-free context (i.e. when the data are observed without measurement errors) and
Recursive kernel density estimators with Laplace errors

whose large and moderate deviation principles were established by Slaoui (2013).
In this paper we suppose that \( f_{0,X}(x) = 0 \), and we let \( \Pi_n = \prod_{j=1}^{n} (1 - \gamma_j) \).
Then the proposed estimator (1) can be rewritten as:

\[
\hat{f}_{n,X}(x) = \Pi_n \sum_{k=1}^{n} \Pi_k^{-1} \gamma_k h_k^{-1} K^{\varepsilon} \left( \frac{x - Y_k}{h_k} \right).
\]  

The aim of this paper is to study the properties of the recursive deconvolution kernel density estimator defined by the stochastic approximation algorithm (1) for Laplace errors, and its comparison with the nonrecursive deconvolution kernel density estimator introduced by Carroll and Hall (1988); Stefanski and Carroll (1990), and defined as

\[
\tilde{f}_{n,X}(x) = \frac{1}{nh_n} \sum_{i=1}^{n} K^{\varepsilon} \left( \frac{x - Y_i}{h_n} \right).
\]

This estimator have been investigated by Carroll and Hall (1988); Stefanski and Carroll (1990); Fan (1991a,b,c, 1992); among many others.

We first compute the bias and the variance of the recursive estimator \( f_{n,X} \) defined by (1). It turns out that they heavily depend on the choice of the stepsize \( (\gamma_n) \). Moreover, we proposed a plug-in estimate which minimize an estimate of the Mean Weighted Integrated Squared Error (MWISE), using the density function as weight function to implement the bandwidth selection of the proposed estimator.

The remainder of the paper is organized as follows. In Section 2, we state our main results. Section 3 is devoted to our application results, first by simulations (Subsection 3.1) and second using real dataset through a plug-in method (Subsection 3.2), we give our conclusion in Section 4, whereas the technical details are deferred to Section 5.
2 Assumptions and main results

We define the following class of regularly varying sequences.

**Definition 1.** Let $\gamma \in \mathbb{R}$ and $(v_n)_{n \geq 1}$ be a nonrandom positive sequence. We say that $(v_n) \in \mathcal{GS}(\gamma)$ if

$$\lim_{n \to +\infty} n \left[ 1 - \frac{v_n-1}{v_n} \right] = \gamma.$$  \hfill (4)

Condition (4) was introduced by Galambos and Seneta (1973) to define regularly varying sequences (see also Bojanic and Seneta (1973), and by Mokkadem and Pelletier (2007) in the context of stochastic approximation algorithms. Noting that the acronym $\mathcal{GS}$ stand for (Galambos & Seneta).

Typical sequences in $\mathcal{GS}(\gamma)$ are, for $b \in \mathbb{R}$, $n^\gamma \log n$, $n^\gamma \log \log n$, and so on.

The assumptions to which we shall refer are the following

(A1) $\varepsilon \sim \mathcal{E}d(\sigma)$, i.e. $f_\varepsilon(x) = \exp (-|x|/\sigma) / (2\sigma)$.

(A2) The function $K$ equal to $K(x) = (2\pi)^{-1/2} \exp (-x^2/2)$.

(A3) $i) \ (\gamma_n) \in \mathcal{GS}(-\alpha)$ with $\alpha \in (1/2, 1]$.
   $ii) \ (h_n) \in \mathcal{GS}(-a)$ with $a \in (0, 1)$.
   $iii) \ \lim_{n \to \infty} (n\gamma_n) \in (\min \{2a, (\alpha - 5a)/2 \}, \infty]$.

(A4) $f_X$ is bounded, differentiable, and $f_X^{(2)}$ is bounded.

Throughout this paper we shall use the following notations:

$$\xi = \lim_{n \to \infty} (n\gamma_n)^{-1},$$  \hfill (5)

$$I_1 = \int_{\mathbb{R}} f_X^2(x) \, dx, \quad I_2 = \int_{\mathbb{R}} \left( f_X^{(2)}(x) \right)^2 f_Y(x) \, dx,$$

$$R(K) = \int_{\mathbb{R}} K^2(z) \, dz, \quad \mu_j(K) = \int_{\mathbb{R}} z^j K(z) \, dz.$$
Our first result is the following proposition, which gives the bias and the variance of $f_{n,X}$.

**Proposition 1** (Bias and variance of $f_{n,X}$). Let Assumptions (A1) – (A4) hold, and assume that $f^{(2)}_X$ is continuous at $x$, then we have

1. If $a \in (0, \alpha/9]$, then
   \[
   \mathbb{E}[f_{n,X}(x)] - f_X(x) = \frac{1}{2(1-2a\xi)} h_n^2 f^{(2)}_X(x) + o(h_n^2). \tag{6}
   \]
   If $a \in (\alpha/9, 1)$, then
   \[
   \mathbb{E}[f_{n,X}(x)] - f_X(x) = o\left(\sqrt{\gamma_n h_n^{-5}}\right). \tag{7}
   \]

2. If $a \in [\alpha/9, 1)$, then
   \[
   \text{Var}\left[f_{n,X}(x)\right] = \frac{3\sigma^4}{8\sqrt{\pi}} \frac{1}{2-(\alpha-5a)\xi} \gamma_n h_n^5 f_Y(x) + o\left(\frac{\gamma_n}{h_n^5}\right) \tag{7}
   \]
   If $a \in (0, \alpha/9)$, then
   \[
   \text{Var}\left[f_{n,X}(x)\right] = o(h_n^4). \tag{7}
   \]

3. If $\lim_{n \to \infty} (n\gamma_n) > \max\{2a, (\alpha-5a)/2\}$, then (6) and (7) hold simultaneously.

The bias and the variance of the estimator $f_{n,X}$ defined by the stochastic approximation algorithm (2) then heavily depend on the choice of the stepsize ($\gamma_n$). Let us now state the following theorem, which gives the weak convergence rate of the estimator $f_{n,X}$ defined in (2).

**Theorem 1** (Weak pointwise convergence rate). Let Assumptions (A1) – (A4) hold, and assume that $f^{(2)}_X$ is continuous at $x$.

1. If there exists $c \geq 0$ such that $\gamma_n^{-1} h_n^9 \to c$, then
   \[
   \sqrt{\gamma_n^{-1} h_n^5} (f_{n,X}(x) - f_X(x)) \quad \xrightarrow{D} \quad \mathcal{N}\left(\frac{\sqrt{c}}{2(1-2a\xi)} f^{(2)}_X(x), \frac{3\sigma^4}{8\sqrt{\pi}} \frac{1}{2-(\alpha-5a)\xi} f_Y(x)\right). \tag{7}
   \]
2. If \( \gamma_n^{-1} h_n^9 \rightarrow \infty \), then
\[
\frac{1}{h_n^2} (f_{n,X}(x) - f_X(x)) \overset{D}{\rightarrow} \frac{1}{2(1 - 2a\xi)} f_X^{(2)}(x),
\]
where \( \overset{D}{\rightarrow} \) denotes the convergence in distribution, \( \mathcal{N} \) the Gaussian-distribution and \( \overset{P}{\rightarrow} \) the convergence in probability.

In order to measure the quality of our recursive estimator (2), we use the following quantity,
\[
MWISE[f_{n,X}] = \int_{\mathbb{R}} (\mathbb{E}(f_{n,X}(x)) - f_X(x))^2 f_Y(x) \, dx
+ \int_{\mathbb{R}} \text{Var}(f_{n,X}(x)) f_Y(x) \, dx.
\]
Moreover, in the case \( a = \alpha/9 \), it follows from the proposition 1 that
\[
MWISE[f_{n,X}] \simeq \frac{3\sigma^4}{8\sqrt{\pi}(2 - (\alpha - 5a)\xi)} \gamma_n h_n^{-5} I_1 + \frac{1}{4(1 - 2a\xi)^2} h_n^4 I_2(8)
\]
The first term in (8) can be much larger than the variance component of the integrated mean squared error of an ordinary recursive kernel density estimator Mokkadem et al. (2009a). This is the price paid for not measuring \( \{\varepsilon_i\}_{i=1}^n \) precisely. Corollary 1 gives the \( MWISE \) of the deconvolution recursive kernel estimators (1) using the centred double exponential error distribution \( f_{\varepsilon}(x) = \exp(-|x|/\sigma)/(2\sigma) \). Throughout this paper, we used standard normal kernel. The following corollary gives the bandwidth which minimize the \( MWISE \).

**Corollary 1.** Let Assumptions (A1)–(A4) hold. To minimize the \( MWISE \) of \( f_{n,X} \), the stepsize \( (\gamma_n) \) must be chosen in \( GS(-1) \), the bandwidth \( (h_n) \) must equal
\[
\left( \frac{15\sigma^4}{8\sqrt{\pi}} \right)^{1/9} \frac{(1 - 2a\xi)^{2/9}}{(2 - (\alpha - 5a)\xi)^{1/9}} \left\{ \frac{I_1}{I_2} \right\}^{1/9} \gamma_n^{1/9}.
\]
Then, the \( MWISE[f_{n,X}] \)
\[
\simeq \frac{5}{4} \left( \frac{15\sigma^4}{8\sqrt{\pi}} \right)^{4/9} (1 - 2a\xi)^{-10/9} (2 - (\alpha - 5a)\xi)^{-4/9} I_1^{4/9} I_2^{5/9} \gamma_n^{4/9}.
\]
The following corollary shows that, for a special choice of the stepsize \( (\gamma_n) = (\gamma_0 n^{-1}) \), which fulfilled that \( \lim_{n \to \infty} n \gamma_n = \gamma_0 \) and that \( (\gamma_n) \in \mathcal{GS}(-1) \), the optimal value for \( h_n \) depend on \( \gamma_0 \) and then the corresponding \( MWISE \) depend on \( \gamma_0 \).

**Corollary 2.** Let Assumptions (A1)–(A4) hold. To minimize the \( MWISE \) of \( f_{n,X} \), the stepsize \( (\gamma_n) \) must be chosen in \( \mathcal{GS}(-1) \), \( \lim_{n \to \infty} n \gamma_n = \gamma_0 \), and the bandwidth \( (h_n) \) must equal

\[
\left( \left( \frac{15 \sigma^4}{16 \sqrt{\pi}} \right)^{1/9} (\gamma_0 - 2/9)^{1/9} \left\{ \frac{I_1}{I_2} \right\}^{1/9} n^{-1/9} \right).
\]

Then, the \( MWISE \) is

\[
MWISE [f_{n,X}] \approx \frac{9}{20} \left( \frac{15 \sigma^4}{16 \sqrt{\pi}} \right)^{4/9} \frac{\gamma_0^2}{(\gamma_0 - 2/9)^{16/9}} I_1^{4/9} I_2^{5/9} n^{-4/9}.
\]

Moreover, the minimum of \( \gamma_0^2 (\gamma_0 - 2/9)^{-16/9} \) is reached at \( \gamma_0 = 1 \); then the bandwidth \( (h_n) \) must equal

\[
\left( 0.906 \sigma^{4/9} \left\{ \frac{I_1}{I_2} \right\}^{1/9} n^{-1/9} \right).
\]

Then, the \( MWISE \) is

\[
MWISE [f_{n,X}] \approx 0.530 \sigma^{16/9} I_1^{4/9} I_2^{5/9} n^{-4/9}.
\]

In order to estimate the optimal bandwidth (9), we must estimate \( I_1 \) and \( I_2 \). We followed the approach of Altman and Leger (1995), which is called the plug-in estimate, and we use the following kernel estimator of \( I_1 \) introduced in Slaoui (2014a) to implement the bandwidth selection in recursive kernel estimator of probability density function in the error-free context and in Slaoui (2014b) to implement the bandwidth selection in recursive kernel estimator of distribution function also in the error-free data context:

\[
\hat{I}_1 = \frac{\Pi_n}{n} \sum_{i,k=1}^{n} \Pi_{i,k}^{-1} \alpha_k b_k^{-1} K_\theta \left( \frac{Y_i - Y_k}{b_k} \right),
\]

(11)
where $K_{\varepsilon}^2$ is a deconvoluting kernel and $b$ is the associated bandwidth, called the pilot bandwidth and $\alpha$ the pilot stepsize. In practice, we take

$$b_n = n^{-\beta} \min \left\{ \hat{s}, \frac{Q_3-Q_1}{1.349} \right\}, \quad \beta \in (0,1)$$ (12)

(see Silverman (1986)) with $\hat{s}$ the sample standard deviation, and $Q_1$, $Q_3$ denoting the first and third quartiles, respectively. In order to estimate $I_1$, we need to estimate the optimal pilot bandwidth and the optimal pilot stepsize. For this purpose, we should calculate the bias and variance of $\hat{I}_1$, we followed the same steps as in Slaoui (2014a) and we showed that in order to minimize the $MISE$ of $\hat{I}_1$, the pilot bandwidth ($b_n$) should belong to $G\mathcal{S} (-2/9)$, and the pilot stepsize ($\alpha_n$) should be equal to $(1.93 n^{-1})$. Then to estimate $I_1$, we use $\hat{I}_1$, with $b_n$ equal to (12), and $\beta = 2/9$ and $(\alpha_n) = (1.93 n^{-1})$.

Furthermore, to estimate $I_2$, we followed the approach of Slaoui (2014a) and we introduced the following kernel estimator:

$$\hat{I}_2 = \frac{\Pi^2_n}{n} \sum_{i,j,k=1}^{n} \Pi_j^{-1} \Pi_k^{-1} \alpha_j' \alpha_k' b_j' b_k' K_{\varepsilon}^{2} \left( \frac{Y_i - Y_j}{b_j'} \right) K_{\varepsilon}^{2} \left( \frac{Y_i - Y_k}{b_k'} \right),$$ (13)

where $K_{\varepsilon}^{2}$ is the second order derivative of a deconvoluting kernel $K_{\varepsilon}$, and $b'$ the associated bandwidth and $\alpha'$ the pilot stepsize. In order to estimate $I_2$, we need to estimate the optimal pilot bandwidth $b_n'$ and the optimal pilot stepsize $\alpha_n'$. For this purpose, we should calculate the bias and variance of $\hat{I}_2$, we followed the same steps as in Slaoui (2014a) and we showed that in order to minimize the $MISE$ of $\hat{I}_2$, the pilot bandwidth ($b_n'$) should belong to $G\mathcal{S} (-3/22)$, and the pilot stepsize ($\alpha_n'$) should be equal to $(1.65 n^{-1})$. Then to estimate $I_2$, we use $\hat{I}_2$, with $b_n'$ equal to (12), and $\beta = 3/22$ and $(\alpha_n') = (1.65 n^{-1})$.

Finally, the plug-in estimator of the bandwidth ($h_n$) using the recursive
algorithm (2) must equal to

\[
\left(0.906 \sigma^{4/9} \left\{ \frac{\hat{I}_1}{I_2} \right\}^{1/9} n^{-1/9}\right).
\] (14)

Then, it follows from (10) that the MWISE can be estimated by

\[
\text{MWISE} \left[ \hat{f}_{n,X} \right] \simeq 0.530 \sigma^{16/9} \hat{I}_1^{4/9} \hat{I}_2^{5/9} n^{-4/9}.
\]

Now, let us recall that under the assumptions (A1), (A2), (A3) ii) and (A4), the MWISE of the nonrecursive deconvolution kernel density estimator \( \tilde{f}_{n,X} \) (see Stefanski and Carroll (1990)) is given by

\[
\text{MWISE} \left[ \tilde{f}_{n,X} \right] \simeq \frac{3\sigma^4}{8\sqrt{\pi} n h_n^5} I_1 + \frac{1}{4} h_n^4 I_2.
\]

Lemma 1 gives the MWISE of the deconvolution nonrecursive kernel density (3) estimator using the centred double exponential error distribution.

**Lemma 1.** Let Assumptions (A1), (A2), (A3) ii) and (A4) hold. To minimize the MWISE of \( \tilde{f}_{n,X} \), the bandwidth \( (h_n) \) must equal

\[
\left(1.006 \sigma^{4/9} \left\{ \frac{I_1}{I_2} \right\}^{1/9} n^{-1/9}\right).
\] (15)

Then, the MWISE is

\[
\text{MWISE} \left[ \tilde{f}_{n,X} \right] \simeq 0.461 \sigma^{16/9} \hat{I}_1^{4/9} \hat{I}_2^{5/9} n^{-4/9}.
\] (16)

To estimate the optimal bandwidth (15), we must estimate \( I_1 \) and \( I_2 \). As suggested by Hall and Maron (1987), we use the following two kernel estimators to estimate respectively \( I_1 \) and \( I_2 \):

\[
\hat{I}_1 = \frac{1}{n (n - 1) b_n} \sum_{i,j=1}^{n} \sum_{i \neq j} K^{(2)} b_n \left( \frac{Y_i - Y_j}{b_n} \right),
\] (17)

\[
\hat{I}_2 = \frac{1}{n^3 b_n^6} \sum_{i,j,k=1}^{n} \sum_{j \neq k} K^{(2)} b_n \left( \frac{Y_i - Y_j}{b_n} \right) K^{(2)} b_n \left( \frac{Y_i - Y_k}{b_n} \right).
\] (18)
The following Lemma gives the bias and variance of $\tilde{I}_2$. We showed that in order to minimize the $MISE$ of $\tilde{I}_1$ respectively of $\tilde{I}_2$, the pilot bandwidth $(b_n)$ respectively $(b'_n)$ must belong to $\mathcal{G}_S(-2/9)$, respectively to $\mathcal{G}_S(-3/22)$. Then, the plug-in estimator of the bandwidth $(h_n)$ using the nonrecursive algorithm (3) must equal to

$$
(1.006 \sigma^{4/9} \left\{ \frac{\tilde{I}_1}{\tilde{I}_2} \right\}^{1/9} n^{-1/9}).
$$

(19)

Then, it follows from (16) that the $MWISE$ can be estimated by

$$
\text{MWISE} \left[ \tilde{f}_{n,X} \right] \simeq 0.461 \sigma^{16/9} \tilde{I}_1^{4/9} \tilde{I}_2^{-5/9} n^{-4/9}.
$$

The following Theorem gives the conditions under which the expected $MWISE$ of the recursive estimator $f_{n,X}$ will be smaller than the expected $MWISE$ of the nonrecursive estimator $\tilde{f}_{n,X}$.

**Theorem 2.** Let the assumptions (A1)-(A2) hold, and the bandwidth $(h_n)$ equal to (19) and the stepsize $(\gamma_n) = (n^{-1})$. We have

$$
\frac{\mathbb{E} \left[ \text{MWISE} \left[ f_{n,X} \right] \right]}{\mathbb{E} \left[ \text{MWISE} \left[ \tilde{f}_{n,X} \right] \right]} < 1 \quad \text{for small sample setting and small error variance}
$$

Then, the expected $MWISE$ of the recursive estimator defined by (2) is smaller than the expected $MWISE$ of the nonrecursive estimator defined by (3) for small sample setting and when the error variance is controlled by the noise to signal ratio.

Following similar step as in Slaoui (2014a), we can proof Theorem 2.

### 3 Applications

The aim of our applications is to compare the performance of the nonrecursive deconvolution density kernel estimator defined in (3) with that of the recursive deconvolution density kernel estimators defined in (1).
3.1 Simulations

The aim of our simulation study is to compare the performance of the nonrecursive estimator defined in (3) with that of the recursive estimators defined in (2).

When applying $f_{n,X}$ one need to choose three quantities:

- The function $K$, we choose the standard normal kernel.
- The stepsize $(\gamma_n) = (\lfloor 5/9 + c \rfloor n^{-1})$, with $c \in [0, 7/9]$.
- The bandwidth $(h_n)$ is chosen to be equal to (14). To estimate $I_1$, we use the estimator $\hat{I}_1$ given in (11), with $K_{\varepsilon}$ is the standard normal kernel, the pilot bandwidth $(b_n)$ is chosen to be equal to (12), with $\beta = 2/9$, and the pilot stepsize $(\alpha_n) = (1.93 n^{-1})$. Moreover, to estimate $I_2$, we use the estimator $\hat{I}_2$ given in (13), with $K_{\varepsilon}$ is the standard normal kernel, the pilot bandwidth $(b_n)$ is chosen to be equal to (12), with $\beta = 2/9$, and the pilot stepsize $(\alpha_n) = (1.93 n^{-1})$.

When applying $\tilde{f}_n$ one need to choose two quantities:

- The function $K$, as in the recursive framework, we use the normal kernel.
- The bandwidth $(h_n)$ is chosen to be equal to (19). To estimate $I_1$, we used the estimator $\tilde{I}_1$ given in (17), with $K_{\varepsilon}$ is the standard normal kernel, the pilot bandwidth $(b_n)$ is chosen to be equal to (12), with $\beta = 2/9$. Moreover, to estimate $I_2$, we used the estimator $\tilde{I}_2$ given in (18), with $K_{\varepsilon}$ is the standard normal kernel, the pilot bandwidth $(b_n)$ is chosen to be equal to (12), with $\beta = 3/22$.

In order to investigate the comparison between the two estimators, we consider $\varepsilon \sim \mathcal{Ed}(\sigma)$ (i.e. centred double exponentielle with the scale parameter...
\( \sigma \). The error variance was controlled by the noise to signal ratio, denoted by \( \text{NSR} \) and defined by \( \text{NSR} = \text{Var}(\varepsilon) / \text{Var}(X) \). We consider three sample sizes: \( n = 50, n = 100 \) and 500, and four density functions: the normal \( \mathcal{N}(0, 2) \) distribution (see Table 1), the chi-squared distribution with four degrees of freedom distribution \( \chi^2(4) \) (see Table 2), the gamma mixture distribution \( 0.4 \mathcal{G}(5) + 0.6 \mathcal{G}(13) \) (see Table 3). For each of these four cases, 500 samples of sizes \( n = 50, n = 100 \) and 500 were generated. For each fixed \( \text{NSR} \in [5\%, 30\%] \), the number of simulations is 500. We denote by \( f^*_i \) the reference density, and by \( f_i \) the test density, and then we compute the following measures: Mean squared Error \( (MSE = n^{-1} \sum_i (f_i - f^*_i)^2) \) and the linear Correlation \( (\text{Cor} = \text{Cov}(f_i, f^*_i) \sigma(f^*_i)^{-1} \sigma(f_i)^{-1}) \).

Figure 1: Qualitative comparison between the nonrecursive estimator (3) and the proposed estimator (1) with the choice of the stepsize \( (\gamma_n) = (n^{-1}) \), for 500 samples of size 200, with \( \text{NSR} \) equal respectively to 5\% (in the top left panel), equal to 10\% (in the top right panel), equal to 20\% (in the down left panel) and equal to 30\% (in the down right panel) for the normal distribution \( X \sim \mathcal{N}(0, 2) \).
Recursive kernel density estimators with Laplace errors

Figure 2: Qualitative comparison between the nonrecursive estimator (3) and the proposed estimator (1) with the choice of the stepsize ($\gamma_n = (n^{-1})$), for 500 samples of size 200, with NSR equal respectively to 5% (in the top left panel), equal to 10% (in the top right panel), equal to 20% (in the down left panel) and equal to 30% (in the down right panel) for the chi-squared distribution with four degrees of freedom distribution $X \sim \chi(4)$.

From Tables 1, 2 and 3, we conclude that

(i) in all the cases of the Table 1, the $MSE$ of the proposed density estimator (1), with the choice of the stepsize ($\gamma_n = (n^{-1})$) is smaller than the nonrecursive deconvolution density estimator (3).

(ii) in all the cases of the Table 2, the $MSE$ of the proposed density estimator (1), with the choice of the stepsize ($\gamma_n = (n^{-1})$) is smaller than the nonrecursive deconvolution density estimator (3), when the NSR equal to 5% or 10%.

(iii) in all the cases of the Table 3, the $MSE$ of the proposed density estimator (1), with the choice of the stepsize ($\gamma_n = (n^{-1})$) is smaller than the nonrecursive deconvolution density estimator (3), when the...
Figure 3: Qualitative comparison between the nonrecursive estimator (3) and the proposed estimator (1) with the choice of the stepsize \((\gamma_n) = (n^{-1})\), for 500 samples of size 200, with NSR equal respectively to 5\% (in the top left panel), equal to 10\% (in the top right panel), equal to 20\% (in the down left panel) and equal to 30\% (in the down right panel) for the gamma mixture distribution \(X \sim 0.4G(5) + 0.6G(13)\).

NSR equal to 5\% and \(n = 50\) or \(n = 100\).

(iv) the MSE decrease as the sample size increase.

(v) the MSE increase as the value of NSR increase.

(vi) the Cor increase as the sample size increase.

(vii) the Cor decrease as the value of NSR increase.

3.2 Real dataset

We use Salvister data which appears in the R package kerdiest (Quintela-del-Riño and Estévez-Pérez (2012)). These data contains the annual peak instantaneous flow levels of the Salt River near Roosevelt, AZ, USA, for the
period 1924-2009, obtained from the National Water Information System. In order to investigate the comparison between the two estimators, we consider the annual peak for 500 samples of Laplacian errors $\varepsilon \sim \mathcal{E}(\sigma)$, with $\text{NSR} \in [5\%, 30\%]$. For each fixed $\text{NSR}$, we computed the mean (over the 500 samples) of $I_1, I_2, h_n$ and $MWISE$. The plug-in estimators (14), (19) requires two kernels to estimate $I_1$ and $I_2$. In both cases we use the normal kernel with $b_n$ and $b'_n$ are given in (12), with $\beta$ equal respectively to $2/9$ and $3/22$.

Figure 4: Qualitative comparison between the nonrecursive estimator (3) and the proposed estimator (1) with the choice of the stepsize $(\gamma_n) = (n^{-1})$, for 500 samples of Laplacian errors with $\text{NSR}$ equal respectively to $5\%$ (in the top left panel), equal to $10\%$ (in the top right panel), equal to $20\%$ (in the down left panel) and equal to $30\%$ (in the down right panel) for the Salvister data of the package kerdiest and through a plug-in method.

From the table 4, we conclude that, the $\hat{MWISE}$ of proposed estimator is quite similar to the $\tilde{MWISE}$ of the nonrecursive estimator. From the Figures 1, 2, 3 and 4, we conclude that the two estimators present a quite similar behavior for all the fixed $\text{NSR}$. 
4 Conclusion

This paper propose an automatic selection of the bandwidth of a probability density function in the case of deconvolution recursive kernel estimators. The estimators are compared to the nonrecursive deconvolution density estimator (3). We showed that using the selected bandwidth and the stepsizes \((\gamma_n) = \left(\frac{n}{2}\right)^{-1}\), the recursive estimator will be better than the nonrecursive one for small sample setting and when the error variance is controlled by the noise to signal ratio. The simulation study corroborated these theoretical results. Moreover, the simulation results indicate that the proposed recursive estimator was more computing efficiency than the nonrecursive estimator.

In conclusion, the proposed estimators allowed us to obtain quite better results than the nonrecursive estimator proposed by Carroll and Hall (1988); Stefanski and Carroll (1990). Moreover, we plan to make an extensions of our method in future and to consider the case of a regression function (see Mokkadem et al. (2009b) and Slaoui (2015a,b,c, 2016)) in the error-free context.

5 Technical proofs

Throughout this section we use the following notation:

\[ Z_n (x) = h_n^{-1} K \left( \frac{x - X_n}{h_n} \right), \]

and

\[ Z_n^\varepsilon (x) = h_n^{-1} K^\varepsilon \left( \frac{x - Y_n}{h_n} \right). \]  \( \text{(20)} \)

Let us first state the following technical lemma.
Lemma 2. Let \((v_n) \in \mathcal{GS}(v^*), (\gamma_n) \in \mathcal{GS}(-\alpha), \) and \(m > 0\) such that \(m - v^*\xi > 0\) where \(\xi\) is defined in (5). We have

\[
\lim_{n \to +\infty} v_n \Pi_n^m \sum_{k=1}^{n} \Pi_k^{-m} \frac{\gamma_k}{v_k} = \frac{1}{m - v^*\xi}.
\]

Moreover, for all positive sequence \((\alpha_n)\) such that \(\lim_{n \to +\infty} \alpha_n = 0\), and all \(\delta \in \mathbb{R}\),

\[
\lim_{n \to +\infty} v_n \Pi_n^m \left[ \sum_{k=1}^{n} \Pi_k^{-m} \frac{\gamma_k}{v_k} \alpha_k + \delta \right] = 0.
\]

Lemma 2 is widely applied throughout the proofs. Let us underline that it is its application, which requires Assumption \((A3)(iii)\) on the limit of \((n\gamma_n)\) as \(n\) goes to infinity.

Our proofs are organized as follows. Proposition 1 in Section 5.1, Theorem 1 in Section 5.2.

5.1 Proof of Proposition 1

Proof. In view of (2) and (20), we have

\[
f_{n,X}(x) - f_X(x) = \Pi_n \sum_{k=1}^{n} \Pi_k^{-1} \gamma_k (Z_k(x) - f_X(x)) + \Pi_n \left( f_{0,X}(x) - f_X(x) \right).
\]

Then, it follows that

\[
\mathbb{E}(f_{n,X}(x)) - f_X(x) = \Pi_n \sum_{k=1}^{n} \Pi_k^{-1} \gamma_k (\mathbb{E}(Z_k(x)) - f_X(x)) + \Pi_n \left( f_{0,X}(x) - f_X(x) \right).
\]

Moreover, an interchange of expectation and integration, justified by Fubini’s Theorem and assumptions \((A1)\) and \((A2)\), shows that \(\mathbb{E}\{Z_k(x) | X_k\} = Z_k(x)\), which ensure that \(\mathbb{E}[Z_k(x)] = \mathbb{E}[Z_k(x)]\). Taylor’s expansion with
integral remainder implies that
\[
E[Z_k(x)] - f_X(x) = \int_{\mathbb{R}} K(z) [f_X(x - zh_k) - f_X(x)] \, dz \\
= h_k^2 \mu_2(K) f_X^{(2)}(x) + h_k^2 \delta_k(x)
\]
with
\[
\delta_k(x) = h_k^{-2} \int_{\mathbb{R}} K(z) [f_X(x - zh_k) - f_X(x) - z^2 \frac{h_k^2}{2} f_X^{(2)}(x)] \, dz,
\]
and, since \(f_X^{(2)}\) is bounded and continuous at \(x\), we have \(\lim_{k \to \infty} \delta_k(x) = 0\).

In the case \(a \leq \alpha/9\), we have \(\lim_{n \to \infty} (n \gamma_n) > 2a\); the application of Lemma 2 then gives
\[
E[f_n,X(x)] - f_X(x) = \frac{1}{2} \mu_2(K) f_X^{(2)}(x) \Pi_n \sum_{k=1}^{n} \Pi_k^{-1} \gamma_k h_k^2 [1 + o(1)] + \Pi_n (f_{0,X}(x) - f_X(x))
\]
\[
= \frac{1}{2} \mu_2(K) f_X^{(2)}(x) \left[ h_n^2 + o(1) \right],
\]
(6) follows. In the case \(a > \alpha/9\), we have \(h_n^2 = o\left(\sqrt{\gamma_n h_n^{-5}}\right)\); since \(\lim_{n \to \infty} (n \gamma_n) > (\alpha - 5a)/2\), Lemma 2 then ensures that
\[
E[f_n,X(x)] - f_X(x) = \Pi_n \sum_{k=1}^{n} \Pi_k^{-1} \gamma_k o\left(\sqrt{\gamma_k h_k^{-5}}\right) + O(\Pi_n)
\]
\[
= o\left(\sqrt{\gamma_n h_n^{-5}}\right),
\]
which gives (7). Now, we have
\[
Var[f_n,X(x)] = \Pi_n \sum_{k=1}^{n} \Pi_k^{-2} \gamma_k^2 Var[Z_k^\varepsilon(x)]
\]
\[
= \Pi_n \sum_{k=1}^{n} \Pi_k^{-2} \gamma_k^2 [f_Y(x) R(K^\varepsilon) + \nu_k(x) - h_k \tilde{\nu}_k(x)]
\]
with \(\nu_k(x) = \int_{\mathbb{R}} (K^\varepsilon(z))^2 [f_Y(x - zh_k) - f_Y(x)] \, dz\), and
\(\tilde{\nu}_k(x) = \left(\int_{\mathbb{R}} K(z) f_X(x - zh_k) \, dz\right)^2\). In view of (A2) and (A4) i), we have
Recursive kernel density estimators with Laplace errors

\( \lim_{k \to \infty} \nu_k (x) = 0 \) and \( \lim_{k \to \infty} h_k \tilde{\nu}_k (x) = 0 \). Then, we have

\[
\text{Var} [f_{n,X} (x)] = \frac{1}{2\pi} \sum_{k=1}^{n} \Pi_k^{-2} \gamma_k h_k^{-1} R (K^\varepsilon) \left[ f_Y (x) + o(1) \right].
\]

Let us now state the following Lemma:

**Lemma 3.**

\[
R (K^\varepsilon) = \frac{1}{2\sqrt{\pi}} \left( 1 + \left( \frac{\sigma}{h_k} \right)^2 + \frac{3}{4} \left( \frac{\sigma}{h_k} \right)^4 \right).
\]

**Proof of Lemma 3**

*Proof.* First, we have

\[
R (K^\varepsilon) = \frac{1}{2\pi} \int_{R} \phi_K^2 (t) \phi_\varepsilon (t/h_n) \, dt.
\]

Moreover, since \( \varepsilon \) follows Laplace errors \( \mathcal{L} (0, \sigma) \), we have \( \phi_\varepsilon (t) = \frac{1}{(1 + \sigma^2 t^2)} \), and since \( K \sim \mathcal{N} (0, 1) \), we have \( \phi_K (t) = \exp (-t^2 / 2) \). Then, it follows that from (22), that

\[
R (K^\varepsilon) = \frac{1}{2\pi} \left\{ \int_{R} \exp (-t^2) \, dt + 2h_n^{-2} \sigma^2 \int_{R} t^2 \exp (-t^2) \, dt + h_n^{-4} \sigma^4 \int_{R} t^4 \exp (-t^2) \, dt \right\}
\]

Moreover, we have \( \int_{R} \exp (-t^2) \, dt = \sqrt{\pi}, \int_{R} t^2 \exp (-t^2) \, dt = \sqrt{\pi} / 4 \) and \( \int_{R} t^4 \exp (-t^2) \, dt = 3\sqrt{\pi} / 8 \), which concludes the proof of Lemma 3.

Now, when \( a \geq \alpha/9 \), we have \( \lim_{n \to \infty} (n\gamma_n) > (\alpha - 5a) / 2 \), and the application of Lemma 2 gives

\[
\text{Var} [f_{n,X} (x)] = \frac{3\sigma^4}{8\sqrt{\pi} 2\pi (2 - (\alpha - 5a) \xi)} \frac{\gamma_n}{h_n^5} \left[ f_Y (x) + o(1) \right]
\]

which proves (7). In the case \( a < \alpha/9 \), we have \( \gamma_n h_n^{-5} = o (h_n^4) \); since \( \lim_{n \to \infty} (n\gamma_n) > 2a \), Lemma 2 then ensures that \( \text{Var} [f_{n,X} (x)] = o (h_n^4) \), which gives (8).
5.2 Proof of Theorem 1

Proof. Let us at first assume that, if $a \geq \alpha/9$, then

$$\sqrt{\gamma_n^{-1} h_n^5} (f_{n,X} (x) - \mathbb{E} [f_{n,X} (x)]) \xrightarrow{D} \mathcal{N} \left( 0, \frac{3 \sigma^4}{8 \sqrt{\pi}} \frac{1}{2 \sqrt{\pi} (2 - (\alpha - 5a) \xi)} f_Y (x) \right), \quad (23)$$

In the case when $a > \alpha/9$, Part 1 of Theorem 1 follows from the combination of (7) and (23). In the case when $a = \alpha/9$, Parts 1 and 2 of Theorem 1 follow from the combination of (6) and (23). In the case $a < \alpha/5$, we have

$$h_n^{-2} (f_{n,X} (x) - f_X (x)) = \frac{1}{\sqrt{\gamma_n^{-1} h_n^9}} \sqrt{\gamma_n^{-1} h_n^5} (f_{n,X} (x) - \mathbb{E} [f_{n,X} (x)])$$

$$+ h_n^{-2} (\mathbb{E} [f_{n,X} (x)] - f_X (x))$$

and the application of (6) gives Part 2 of Theorem 1.

We now prove (23). In view of (2), we have

$$f_{n,X} (x) - \mathbb{E} [f_{n,X} (x)]$$

$$= (1 - \gamma_n) (f_{n-1,X} (x) - \mathbb{E} [f_{n-1,X} (x)]) + \gamma_n (Z_n^x (x) - \mathbb{E} [Z_n^x (x)])$$

$$= \Pi_n \sum_{k=1}^n \Pi_k^{-1} \gamma_k (Z_k^x (x) - \mathbb{E} [Z_k^x (x)]) .$$

Set

$$Y_k (x) = \Pi_k^{-1} \gamma_k (Z_k^x (x) - \mathbb{E} (Z_k^x (x))) .$$
The application of Lemma 2 ensures that

\[ v_n^2 = \sum_{k=1}^{n} \text{Var}(Y_k(x)) \]

\[ = \sum_{k=1}^{n} \Pi_k^{-2} \gamma_k^2 \text{Var}(Z_k^\varepsilon(x)) \]

\[ = \sum_{k=1}^{n} \Pi_k^{-2} \gamma_k^2 \frac{1}{h_k} [f_Y(x) R(K^\varepsilon) + o(1)] \]

\[ = \frac{3\sigma^4}{8\sqrt{\pi}} \frac{1}{2\sqrt{\pi} (2 - (\alpha - 5a) \xi)} \frac{1}{n^2 h_n^2} \left[ f_Y(x) + o(1) \right]. \]

On the other hand, we have, for all \( p > 0 \),

\[ \mathbb{E} \left[ |Z_k^\varepsilon(x)|^{2+p} \right] = O \left( \frac{1}{h_k^{(1+p)}} \right), \]

and, since \( \lim_{n \to \infty} (n\gamma_n) > (\alpha - 5a)/2 \), there exists \( p > 0 \) such that \( \lim_{n \to \infty} (n\gamma_n) > \frac{1+p}{2+p} (\alpha - 5a) \). Applying Lemma 2, we get

\[ \sum_{k=1}^{n} \mathbb{E} \left[ |Y_k(x)|^{2+p} \right] = O \left( \sum_{k=1}^{n} \Pi_k^{-2-p} \gamma_k^{2+p} \mathbb{E} \left[ |Z_k^\varepsilon(x)|^{2+p} \right] \right) \]

\[ = O \left( \sum_{k=1}^{n} \Pi_k^{-2-p} \gamma_k^{2+p} \frac{1}{h_k^{(1+p)}} \right) \]

\[ = O \left( \frac{\gamma_1^{1+p}}{\Pi_n^{2+p} h_n^{(1+p)}} \right), \]

and we thus obtain

\[ \frac{1}{v_n^{2+p}} \sum_{k=1}^{n} \mathbb{E} \left[ |Y_k(x)|^{2+p} \right] = O \left( \frac{\gamma_n^{p/2} h_n^{4+3/2p}}{h_n^{(1+p)}} \right) = o(1). \]

The convergence in (23) then follows from the application of Lyapounov’s Theorem. \( \Box \)
References


Yousri Slaoui
Associate professor at the University of Poitiers
Laboratoire de Mathmatiques et Applications
UMR 7348 du CNRS, Téléport 2- BP 30179
11 Boulevard Marie et Pierre Curie
86962 Futuroscope Chasseneuil, France
E-mail: Yousri.Slaoui@math.univ-poitiers.fr
<table>
<thead>
<tr>
<th>$n$</th>
<th>nonrecursive</th>
<th>estimator 1</th>
<th>estimator 2</th>
<th>estimator 3</th>
<th>estimator 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>4.18e$^{-05}$</td>
<td>1.20e$^{-04}$</td>
<td>3.68e$^{-05}$</td>
<td>2.36e$^{-05}$</td>
<td>2.11e$^{-05}$</td>
</tr>
<tr>
<td></td>
<td>Cor</td>
<td>0.99833</td>
<td>0.99887</td>
<td>0.99899</td>
<td>0.99904</td>
</tr>
<tr>
<td>100</td>
<td>3.12e$^{-05}$</td>
<td>6.90e$^{-05}$</td>
<td>2.38e$^{-05}$</td>
<td>1.80e$^{-05}$</td>
<td>1.79e$^{-05}$</td>
</tr>
<tr>
<td></td>
<td>Cor</td>
<td>0.99864</td>
<td>0.99913</td>
<td>0.99918</td>
<td>0.99916</td>
</tr>
<tr>
<td>500</td>
<td>2.13e$^{-05}$</td>
<td>2.93e$^{-05}$</td>
<td>1.51e$^{-05}$</td>
<td>1.32e$^{-05}$</td>
<td>1.29e$^{-05}$</td>
</tr>
<tr>
<td></td>
<td>Cor</td>
<td>0.99913</td>
<td>0.99937</td>
<td>0.99944</td>
<td>0.99946</td>
</tr>
<tr>
<td>50</td>
<td>6.89e$^{-05}$</td>
<td>1.70e$^{-04}$</td>
<td>7.10e$^{-05}$</td>
<td>5.36e$^{-05}$</td>
<td>5.13e$^{-05}$</td>
</tr>
<tr>
<td></td>
<td>Cor</td>
<td>0.99727</td>
<td>0.99780</td>
<td>0.99790</td>
<td>0.99789</td>
</tr>
<tr>
<td>100</td>
<td>6.61e$^{-05}$</td>
<td>1.26e$^{-04}$</td>
<td>6.45e$^{-05}$</td>
<td>5.35e$^{-05}$</td>
<td>5.11e$^{-05}$</td>
</tr>
<tr>
<td></td>
<td>Cor</td>
<td>0.99725</td>
<td>0.99753</td>
<td>0.99771</td>
<td>0.99777</td>
</tr>
<tr>
<td>500</td>
<td>4.06e$^{-05}$</td>
<td>5.80e$^{-05}$</td>
<td>3.71e$^{-05}$</td>
<td>3.35e$^{-05}$</td>
<td>3.25e$^{-05}$</td>
</tr>
<tr>
<td></td>
<td>Cor</td>
<td>0.99844</td>
<td>0.99847</td>
<td>0.99865</td>
<td>0.99873</td>
</tr>
<tr>
<td>50</td>
<td>1.14e$^{-03}$</td>
<td>2.44e$^{-03}$</td>
<td>1.31e$^{-03}$</td>
<td>1.10e$^{-03}$</td>
<td>1.09e$^{-03}$</td>
</tr>
<tr>
<td></td>
<td>Cor</td>
<td>0.99586</td>
<td>0.99615</td>
<td>0.99616</td>
<td>0.99601</td>
</tr>
<tr>
<td>100</td>
<td>1.05e$^{-04}$</td>
<td>1.93e$^{-04}$</td>
<td>1.17e$^{-04}$</td>
<td>1.02e$^{-04}$</td>
<td>9.98e$^{-05}$</td>
</tr>
<tr>
<td></td>
<td>Cor</td>
<td>0.99591</td>
<td>0.99564</td>
<td>0.99595</td>
<td>0.99604</td>
</tr>
<tr>
<td>500</td>
<td>8.48e$^{-05}$</td>
<td>1.17e$^{-04}$</td>
<td>8.82e$^{-05}$</td>
<td>8.33e$^{-05}$</td>
<td>8.27e$^{-05}$</td>
</tr>
<tr>
<td></td>
<td>Cor</td>
<td>0.99692</td>
<td>0.99665</td>
<td>0.99691</td>
<td>0.99698</td>
</tr>
</tbody>
</table>

Table 1: Quantitative comparison between the nonrecursive estimator (3) and four estimators; estimator 1 correspond to the estimator (1) with the choice of $(\gamma_n) = ([5/9] n^{-1})$, estimator 2 correspond to the estimator (1) with the choice of $(\gamma_n) = ([7/9] n^{-1})$, estimator 3 correspond to the estimator (1) with the choice of $(\gamma_n) = (n^{-1})$ and estimator 4 correspond to the estimator (1) with the choice of $(\gamma_n) = ([4/3] n^{-1})$. Here we consider the normal distribution $X \sim \mathcal{N}(0, 2)$ with $\text{NSR} = 5\%$ in the first block, $\text{NSR} = 10\%$ in the second block and $\text{NSR} = 20\%$ in the last block, we consider three sample sizes $n = 50$, $n = 100$ and $n = 500$, the number of simulations is 500, and we compute the Mean squared error ($MSE$) and the linear correlation ($Cor$).
Table 2: Quantitative comparison between the nonrecursive estimator (3) and four estimators; estimator 1 correspond to the estimator (1) with the choice of \( (\gamma_n) = \left(\frac{5}{9}\right)n^{-1} \), estimator 2 correspond to the estimator (1) with the choice of \( (\gamma_n) = \left(\frac{7}{9}\right)n^{-1} \), estimator 3 correspond to the estimator (1) with the choice of \( (\gamma_n) = \left(n^{-1}\right) \) and estimator 4 correspond to the estimator (1) with the choice of \( (\gamma_n) = \left(\frac{4}{3}\right)n^{-1} \). Here we consider the chi-squared distribution with four degrees of freedom distribution \( X \sim \chi^2(4) \) with \( \text{NSR} = 5\% \) in the first block, \( \text{NSR} = 10\% \) in the second block and \( \text{NSR} = 20\% \) in the last block, we consider three sample sizes \( n = 50, n = 100 \) and \( n = 500 \), the number of simulations is 500, and we compute the Mean squared error (\( \text{MSE} \)) and the linear correlation (\( \text{Cor} \)).

<table>
<thead>
<tr>
<th></th>
<th>nonrecursive</th>
<th>estimator 1</th>
<th>estimator 2</th>
<th>estimator 3</th>
<th>estimator 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n = 50 )</td>
<td>( \text{NSR} = 5% )</td>
<td>( \text{NSR} = 10% )</td>
<td>( \text{NSR} = 20% )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{MSE} )</td>
<td>1.07e^{-04}</td>
<td>1.71e^{-04}</td>
<td>1.01e^{-04}</td>
<td>8.58e^{-05}</td>
<td>8.23e^{-05}</td>
</tr>
<tr>
<td>( \text{Cor} )</td>
<td>0.98876</td>
<td>0.99014</td>
<td>0.99075</td>
<td>0.99094</td>
<td>0.99106</td>
</tr>
<tr>
<td>( n = 100 )</td>
<td>( \text{NSR} = 5% )</td>
<td>( \text{NSR} = 10% )</td>
<td>( \text{NSR} = 20% )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{MSE} )</td>
<td>8.80e^{-05}</td>
<td>1.18e^{-04}</td>
<td>7.789e^{-05}</td>
<td>7.12e^{-05}</td>
<td>7.15e^{-05}</td>
</tr>
<tr>
<td>( \text{Cor} )</td>
<td>0.99080</td>
<td>0.99218</td>
<td>0.99246</td>
<td>0.99243</td>
<td>0.99229</td>
</tr>
<tr>
<td>( n = 500 )</td>
<td>( \text{NSR} = 5% )</td>
<td>( \text{NSR} = 10% )</td>
<td>( \text{NSR} = 20% )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{MSE} )</td>
<td>6.16e^{-05}</td>
<td>6.96e^{-05}</td>
<td>5.32e^{-05}</td>
<td>5.03e^{-05}</td>
<td>4.97e^{-05}</td>
</tr>
<tr>
<td>( \text{Cor} )</td>
<td>0.99384</td>
<td>0.99444</td>
<td>0.99483</td>
<td>0.99493</td>
<td>0.99497</td>
</tr>
<tr>
<td>( n = 50 )</td>
<td>( \text{NSR} = 10% )</td>
<td>( \text{NSR} = 20% )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{MSE} )</td>
<td>0.000141</td>
<td>0.000235</td>
<td>0.000136</td>
<td>0.000131</td>
<td></td>
</tr>
<tr>
<td>( \text{Cor} )</td>
<td>0.98860</td>
<td>0.98613</td>
<td>0.98629</td>
<td>0.98640</td>
<td></td>
</tr>
<tr>
<td>( n = 100 )</td>
<td>( \text{NSR} = 10% )</td>
<td>( \text{NSR} = 20% )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{MSE} )</td>
<td>0.000128</td>
<td>0.000186</td>
<td>0.000125</td>
<td>0.000124</td>
<td></td>
</tr>
<tr>
<td>( \text{Cor} )</td>
<td>0.98769</td>
<td>0.98769</td>
<td>0.98772</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( n = 500 )</td>
<td>( \text{NSR} = 10% )</td>
<td>( \text{NSR} = 20% )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{MSE} )</td>
<td>1.00e^{-04}</td>
<td>1.24e^{-04}</td>
<td>1.01e^{-04}</td>
<td>9.78e^{-05}</td>
<td>9.70e^{-05}</td>
</tr>
<tr>
<td>( \text{Cor} )</td>
<td>0.99048</td>
<td>0.99053</td>
<td>0.99071</td>
<td>0.99079</td>
<td></td>
</tr>
<tr>
<td>( n = 50 )</td>
<td>( \text{NSR} = 20% )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{MSE} )</td>
<td>0.000235</td>
<td>0.000380</td>
<td>0.000280</td>
<td>0.000255</td>
<td>0.000249</td>
</tr>
<tr>
<td>( \text{Cor} )</td>
<td>0.97731</td>
<td>0.97515</td>
<td>0.97547</td>
<td>0.97559</td>
<td></td>
</tr>
<tr>
<td>( n = 100 )</td>
<td>( \text{NSR} = 20% )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{MSE} )</td>
<td>0.000190</td>
<td>0.000288</td>
<td>0.000222</td>
<td>0.000208</td>
<td>0.000206</td>
</tr>
<tr>
<td>( \text{Cor} )</td>
<td>0.98223</td>
<td>0.98039</td>
<td>0.98062</td>
<td>0.98066</td>
<td></td>
</tr>
<tr>
<td>( n = 500 )</td>
<td>( \text{NSR} = 20% )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{MSE} )</td>
<td>0.000169</td>
<td>0.000217</td>
<td>0.000188</td>
<td>0.000183</td>
<td>0.000183</td>
</tr>
<tr>
<td>( \text{Cor} )</td>
<td>0.98483</td>
<td>0.98350</td>
<td>0.98358</td>
<td>0.98354</td>
<td></td>
</tr>
</tbody>
</table>
The MSE of simulations is 500, and we compute the Mean squared error (MSE) for three sample sizes, first block, NSR = 5\%, Cor = 10\% in the second block, NSR = 10\%, Cor = 20\% in the last block, NSR = 20\%, Cor = 5\% in the first block, NSR = 5\% in the second block and NSR = 5\% in the last block.

<table>
<thead>
<tr>
<th>n = 50</th>
<th>nonrecursive</th>
<th>estimator 1</th>
<th>estimator 2</th>
<th>estimator 3</th>
<th>estimator 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>1.85e-05</td>
<td>4.04e-05</td>
<td>2.24e-05</td>
<td>1.92e-05</td>
<td>1.83e-05</td>
</tr>
<tr>
<td>Cor</td>
<td>0.99067</td>
<td>0.98930</td>
<td>0.98998</td>
<td>0.99030</td>
<td>0.99063</td>
</tr>
<tr>
<td>n = 100</td>
<td>MSE</td>
<td>1.83e-05</td>
<td>3.19e-05</td>
<td>2.10e-05</td>
<td>1.90e-05</td>
</tr>
<tr>
<td>Cor</td>
<td>0.99044</td>
<td>0.98909</td>
<td>0.98976</td>
<td>0.99005</td>
<td>0.99031</td>
</tr>
<tr>
<td>n = 500</td>
<td>MSE</td>
<td>1.41e-05</td>
<td>1.88e-05</td>
<td>1.53e-05</td>
<td>1.48e-05</td>
</tr>
<tr>
<td>Cor</td>
<td>0.99278</td>
<td>0.99188</td>
<td>0.99232</td>
<td>0.99244</td>
<td>0.99247</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>n = 50</th>
<th>NSR = 5%</th>
<th>nonrecursive</th>
<th>estimator 1</th>
<th>estimator 2</th>
<th>estimator 3</th>
<th>estimator 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>3.06e-05</td>
<td>5.66e-05</td>
<td>3.74e-05</td>
<td>3.42e-05</td>
<td>3.37e-05</td>
<td></td>
</tr>
<tr>
<td>Cor</td>
<td>0.98376</td>
<td>0.98136</td>
<td>0.98174</td>
<td>0.98174</td>
<td>0.98176</td>
<td></td>
</tr>
<tr>
<td>n = 100</td>
<td>MSE</td>
<td>2.84e-05</td>
<td>4.64e-05</td>
<td>3.39e-05</td>
<td>3.17e-05</td>
<td>3.13e-05</td>
</tr>
<tr>
<td>Cor</td>
<td>0.98561</td>
<td>0.98323</td>
<td>0.98374</td>
<td>0.98379</td>
<td>0.98374</td>
<td></td>
</tr>
<tr>
<td>n = 500</td>
<td>MSE</td>
<td>2.74e-05</td>
<td>3.50e-05</td>
<td>3.05e-05</td>
<td>2.98e-05</td>
<td>2.98e-05</td>
</tr>
<tr>
<td>Cor</td>
<td>0.98609</td>
<td>0.98425</td>
<td>0.98470</td>
<td>0.98475</td>
<td>0.98471</td>
<td></td>
</tr>
<tr>
<td>n = 50</td>
<td>NSR = 10%</td>
<td>MSE</td>
<td>5.39e-05</td>
<td>9.02e-05</td>
<td>6.62e-05</td>
<td>6.08e-05</td>
</tr>
<tr>
<td>Cor</td>
<td>0.97311</td>
<td>0.96760</td>
<td>0.96878</td>
<td>0.96936</td>
<td>0.96996</td>
<td></td>
</tr>
<tr>
<td>n = 100</td>
<td>MSE</td>
<td>5.27e-05</td>
<td>7.62e-05</td>
<td>6.16e-05</td>
<td>5.88e-05</td>
<td>5.84e-05</td>
</tr>
<tr>
<td>Cor</td>
<td>0.97295</td>
<td>0.96818</td>
<td>0.96925</td>
<td>0.96959</td>
<td>0.96963</td>
<td></td>
</tr>
<tr>
<td>n = 500</td>
<td>MSE</td>
<td>4.82e-05</td>
<td>6.00e-05</td>
<td>5.40e-05</td>
<td>5.29e-05</td>
<td>5.27e-05</td>
</tr>
<tr>
<td>Cor</td>
<td>0.97603</td>
<td>0.97252</td>
<td>0.97326</td>
<td>0.97342</td>
<td>0.97345</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Quantitative comparison between the nonrecursive estimator (3) and four estimators; estimator 1 correspond to the estimator (1) with the choice of (γ_n) = ([5/9] n^{-1}), estimator 2 correspond to the estimator (1) with the choice of (γ_n) = ([7/9] n^{-1}), estimator 3 correspond to the estimator (1) with the choice of (γ_n) = (n^{-1}) and estimator 4 correspond to the estimator (1) with the choice of (γ_n) = ([4/3] n^{-1}). Here we consider the gamma mixture distribution X ~ 0.4G(5) + 0.6G(13) with NSR = 5\% in the first block, NSR = 10\% in the second block and NSR = 20\% in the last block, we consider three sample sizes n = 50, n = 100 and n = 500, the number of simulations is 500, and we compute the Mean squared error (MSE) and the linear correlation (Cor).
<table>
<thead>
<tr>
<th></th>
<th>$I_1$</th>
<th>$I_2$</th>
<th>$h_n$</th>
<th>$MWISE$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NSR = 5%</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonrecursive</td>
<td>$1.9286e^{-2}$</td>
<td>$1.1327e^{-8}$</td>
<td>6.0943</td>
<td>$7.2029e^{-6}$</td>
</tr>
<tr>
<td>Recursive</td>
<td>$1.8140e^{-2}$</td>
<td>$8.2467e^{-9}$</td>
<td>6.0633</td>
<td>$7.2020e^{-6}$</td>
</tr>
<tr>
<td><strong>NSR = 10%</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonrecursive</td>
<td>$1.9286e^{-2}$</td>
<td>$1.1339e^{-8}$</td>
<td>7.1083</td>
<td>$1.3346e^{-5}$</td>
</tr>
<tr>
<td>Recursive</td>
<td>$1.8143e^{-2}$</td>
<td>$8.2668e^{-9}$</td>
<td>7.0712</td>
<td>$1.3355e^{-5}$</td>
</tr>
<tr>
<td><strong>NSR = 20%</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonrecursive</td>
<td>$1.9291e^{-2}$</td>
<td>$1.1338e^{-8}$</td>
<td>8.2924</td>
<td>$2.4714e^{-5}$</td>
</tr>
<tr>
<td>Recursive</td>
<td>$1.8147e^{-2}$</td>
<td>$8.2666e^{-9}$</td>
<td>8.2490</td>
<td>$2.4733e^{-5}$</td>
</tr>
<tr>
<td><strong>NSR = 30%</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonrecursive</td>
<td>$1.9291e^{-2}$</td>
<td>$1.1342e^{-8}$</td>
<td>9.0739</td>
<td>$3.5446e^{-5}$</td>
</tr>
<tr>
<td>Recursive</td>
<td>$1.8145e^{-2}$</td>
<td>$8.2626e^{-9}$</td>
<td>9.0271</td>
<td>$3.5453e^{-5}$</td>
</tr>
</tbody>
</table>

Table 4: The comparison between the $MWISE$ of the Nadaraya’s distribution estimator (3) and the $MWISE$ of the proposed distribution estimator (1) with the choice of the stepsize $(\gamma_n) = (n^{-1})$ via the Salvister data of the package kerdiest and through a plug-in method, with NSR equal to 5% in the first block, 10% in the second block, 20% in the third block and 30% in the last block the number of simulations is 500.