Abstract: A new method is developed for reducing the boundary bias of kernel density and density derivative estimators. Basic asymptotic properties are derived. Simulations demonstrate the viability and efficacy of the approach compared to several alternatives.

1. Introduction

Bias reduction in kernel estimation has received considerable attention in the statistics literature. As Jones and Foster (1993, 1996) and Foster (1995) survey, most of the suggestions in this regard can be seen as special cases of generalized jackknifing in which linear combinations of kernels are constructed to reduce bias. Several authors have considered bias reduction in the context of a boundary problem. One way of viewing the boundary problem is that the effective support of the kernel becomes truncated so that the kernel neither integrates to one nor do its lower moments vanish as is usually required for bias reduction. Hall and Wehrly (1991) suggest “reflecting” data points around the boundary. This is somewhat ad hoc however and does not necessarily remove bias. Gasser and Muller (1979) have suggested various boundary kernels which mix the kernel with a polynomial constructed so that the mixture has vanishing lower moments. Contrasted with this, say, indirect approach, Rice (1984) suggested a direct method for eliminating the second order bias using a linear combination of two estimators. However, this approach does not readily generalize for higher order bias reduction. Jones (1993) shows that a linear combination of a kernel and its derivative can also remove the second order bias. However, he does not consider how to remove higher order bias.

In some situations one may wish to remove bias to a higher order. This may be the case in purely nonparametric estimation procedures when one wants a higher rate of convergence. Also, in some semiparametric problems, the kernel estimator of the nonparametric component of a model is assumed to have the higher order bias removed. For example, this is the case in Klein and Spady (1993) and Bearse, Canals and Rilstone (2007). In many instances, such as with duration models, boundary problems are the norm rather than the exception.

In this paper we propose an alternative direct approach to higher order bias reduction. In the simulations we have conducted, we find that our approach has distinct advantages over its competitors.

The intuition of our approach is as follows. Let $Y_i, i = 1, \ldots, N$, be i.i.d. random variables whose common density, $f$, has support $[0, \infty)$. It is assumed that $f(0) > 0$. We focus on estimating at points close to zero. Right boundary problems and non-zero boundary problems can be dealt with in an analogous fashion. Let $\hat{f}$ be a standard kernel density estimator:

---

1Corresponding author. Email address: pril@yorku.ca. Funding for Rilstone from the Social Sciences and Humanities Research Council of Canada.
\[ \hat{f}(y) = \frac{1}{N\gamma} \sum_{i=1}^{N} K \left( \frac{y - Y_i}{\gamma} \right) \]  

(1.1)

where \( K \) has support \([-1, 1]\). Let \( g^{(s)} \) denote the \( s \)-order derivative of a function \( g \). Put \( g^{[m]} = (g, g^{(1)}, g^{(2)}, \ldots, g^{(m)})^T \), where superscript \( T \) indicates transposition. Under standard regularity conditions, it is straightforward that the expected value of \( \hat{f}(y) \) can be derived as

\[
\mathcal{E} \left[ \hat{f}(y) \right] = \frac{1}{\gamma} \mathcal{E} \left[ K \left( \frac{y - Y_i}{\gamma} \right) \right] = \int_{-1}^{y/\gamma} K(w) f(y - w\gamma) dw \\
= f(y) \int_{-1}^{y/\gamma} K(w) dw + \gamma f^{(1)}(y) \int_{-1}^{y/\gamma} w K(w) dw \\
+ \cdots + \gamma^s f^{(s)}(y) \int_{-1}^{y/\gamma} K(w) \left( -w \right)^s \frac{1}{s!} dw + \gamma^{s+1} O(1)  
\]

(1.2)

By inspection, the first order bias can be removed by dividing the usual estimator by \( \int_{-1}^{y/\gamma} K(w) dw \). Rice’s (1984) proposal of taking a linear combination of two kernel estimators effectively provides a discrete approximation to \( f^{(1)}(y) \). This approach can be extended to removing higher order bias, but the resulting estimator is somewhat unwieldy.

Our approach is as follows. By inspection, it is clear that any unbiased estimator of \( f^{(1)}(y) \) can be used to remove the bias of \( \hat{f} \) to order \( \gamma^2 \). However the usual kernel density estimator of \( f^{(1)}(y) \) is biased in the same manner that \( \hat{f} \) is. In fact the second order bias of \( \hat{f}^{(1)} \) depends on \( f, f^{(1)} \) and \( f^{(2)} \). More generally, it can be shown that the bias of, say, \( \hat{f}^{(j)}, j \leq s \) depends on \( f, f^{(1)}, \ldots, f^{(s)} \). In Section 2 we show how to construct a linear combination of \( \hat{f} \) and its derivatives to obtain an estimator, unbiased to arbitrary order.

To illustrate this in the second order case we have the following. Put \( K^{[1]} = (K, K^{(1)})^T \). By standard manipulations we have

\[
\mathcal{E} \left[ K^{[1]} \left( \frac{y - Y_i}{\gamma} \right) \right] = f(y) \gamma \int_{-1}^{y/\gamma} K^{[1]}(w) dw - \gamma^2 f^{(1)}(y) \int_{-1}^{y/\gamma} K^{[1]}(w) wdw + \gamma^3 O(1).  
\]

(1.3)

Let

\[ Q_2(y/\gamma) = \int_{-1}^{y/\gamma} (1, -w)K^{[1]}(w) dw, \quad \Gamma_2 = \begin{pmatrix} \gamma & 0 \\ 0 & \gamma_2 \end{pmatrix} \]

(1.4)

so that \( Q_2 \) is a matrix of incomplete moments of \( K^{[1]} \). (Note that, for most kernels used in estimation, this is simply the identity matrix for \( y \geq \gamma \).)

It is straightforward that

\[
\mathcal{E} \left[ K^{[1]} \left( \frac{y - Y_i}{\gamma} \right) \right] = Q_2(y/\gamma) \Gamma_2 f^{[1]}(y) + \gamma^2 O(1).  
\]

(1.5)
Therefore with

\[ \tilde{f}^{[1]}(y) = \Gamma_2^{-1}Q_2(y/\gamma)^{-1} \frac{1}{N} \sum_{i=1}^{N} K^{[1]} \left( \frac{y - Y_i}{\gamma} \right) \]  

we have

\[ \mathcal{E} \left[ \tilde{f}^{[1]}(y) \right] = \Gamma_2^{-1}Q_2(y/\gamma)^{-1}\mathcal{E} \left[ K^{[1]} \left( \frac{y - Y_i}{\gamma} \right) \right] 
= f^{[1]}(y) + \gamma^3 T^{-1} O(1) \]  

so that, using the first element of this to estimate \( f(y) \) we have

\[ \mathcal{E} \left[ \tilde{f}(y) \right] = f(y) + \gamma^2 O(1). \]  

Also note that the second element of \( \tilde{f}^{[1]}(y) \) provides an estimator of \( f^{(1)}(y) \) which is also unbiased to order \( O(\gamma^2) \).

In the next section we show how bias reduction can be done to arbitrary order. We also derive the pointwise variance and hence get a pointwise rate of convergence. In Section 3 we use a simple simulation to show how the procedure works in practice and compare its performance to unadjusted kernels and boundary kernels. Section 4 concludes.

### 2. Asymptotic properties and example

Before stating the estimator, some additional notation is useful. Put

\[ W_s(w) = \left( 1, -w, \frac{(-w)^2}{2!}, \ldots, \frac{(-w)^s}{s!} \right) \]  

and

\[ \Gamma_s = \text{Diag}(\gamma, \gamma^2, \ldots, \gamma^s) \]  

Define an \( s \times s \) matrix of partial moments for \( K^{[s-1]} \) by

\[ Q_s(y/\gamma) = \int_{-1}^{y/\gamma} K^{[s-1]}(w)W_s(w) \, dw \]  

and define an \( 1 \times s \) row vector \( \iota_1 = (1, 0, \ldots, 0) \). The estimator is thus given by

\[ \tilde{f}(y) = \iota_1 \Gamma_s^{-1} Q_s^{-1}(y/\gamma) \frac{1}{N} \sum_{i=1}^{N} K^{[s-1]} \left( \frac{y - Y_i}{\gamma} \right). \]  

We make standard assumptions about the kernel and window width as follows. \( K \) is bounded with support \([-1, 1]\). \( \int K(w)dw = 1. \) \( K(w) \) is \( s \)-times differentiable. \( K(w) \) is an \( s \)-order kernel such that, for some \( s \geq 1, \int w^m K(w)dw = 0 \) for \( m = 1, \ldots, s - 1 \) and \( \int |w|^s |K(w)| \, dw < \infty \). The window width sequence satisfies \( \lim_{N \to \infty} \gamma = 0 \) and \( \lim_{N \to \infty} N \gamma = \infty \). \( Q_s^{-1} \) exists uniformly and is finite.

**Proposition 1.** Suppose that \( f(y) \) is differentiable to order \( s \) and these derivatives are uniformly bounded. Then, uniformly in \( y \geq 0, \)
\[ E \left[ f(y) \right] - f(y) = O(\gamma^s). \]

**Proof:** Using a change of variables and an s'th order Taylor series expansion of \( f \) we have,

\[
E \left[ \mathcal{K}^{(s-1)} \left( \frac{y - Y_i}{\gamma} \right) \right]
= \int_{-1}^{\infty} K^{(s-1)} \left( \frac{y - Y_i}{\gamma} \right) f(Y_i) dY_i
= \gamma \int_{-1}^{y/\gamma} K^{(s-1)} (w) f(y - w\gamma) dw
= \gamma \int_{-1}^{y/\gamma} K^{(s-1)} (w) \left[ \sum_{k=0}^{s-1} \frac{f^{(k)}(y)(-w\gamma)^k}{k!} + \frac{f^{(s)}(\bar{y})(-w\gamma)^s}{s!} \right] dw
= Q_s(y/\gamma) \Gamma_s f^{(s-1)}(y) + \gamma^{s+1} \int_{-1}^{y/\gamma} K^{(s-1)} (w) f^{(s)}(\bar{y})(-w\gamma)^s dw
\tag{2.5}
\]

where \( \bar{y} \) is a mean value. Since \( Q_s^{-1}(y/\gamma) \), \( K^{(s-1)} \) and \( f^{(s)} \) are bounded, we have

\[
\left| E \left[ \mathcal{K}^{(s-1)} \left( \frac{y - Y_i}{\gamma} \right) \right] - f(y) \right| \leq \gamma^s C,
\tag{2.6}
\]

uniformly in \( y \geq 0 \). QED

Put

\[
M_N = Q_s(y/\gamma)^{-1} E \left[ \mathcal{K}^{(s-1)} \left( \frac{y - Y_i}{\gamma} \right) \mathcal{K}^{(s-1)} \left( \frac{y - Y_i}{\gamma} \right)^T \right] (Q_s(y/\gamma)^{-1})^T
= \gamma f(y) Q_s(y/\gamma)^{-1} \int_{-1}^{y/\gamma} K^{(s-1)} (w) K^{(s-1)} (w)^T dw (Q_s(y/\gamma)^{-1})^T + o(\gamma).
\tag{2.7}
\]

The variance of the estimator is given by the following result.

**Proposition 2.** Suppose that \( f(y) \) is differentiable to order \( s \) and these derivatives are uniformly bounded. Then,

\[
\text{Var} \left[ \tilde{f}(y) \right] = \frac{1}{N\gamma} f(y) \int_{-1}^{1} K(w)^2 dw + o \left( \frac{1}{N\gamma} \right)
\]

**Proof:** The result follows by standard change of variables as follows.

\[
\text{Var} \left[ \tilde{f}^{(s-1)}(y) \right]
= \frac{1}{N} \Gamma_s^{-1} Q_s(y/\gamma)^{-1} \text{Var} \left[ \mathcal{K}^{(s-1)} \left( \frac{y - Y_i}{\gamma} \right) \right] (Q_s(y/\gamma)^{-1})^T \Gamma_s^{-1}
= \frac{1}{N} \Gamma_s^{-1} M_N \Gamma_s^{-1} + \Gamma_s^{-1} o \left( \frac{\gamma}{N} \right) \Gamma_s^{-1}.
\tag{2.8}
\]
Note that $\iota_s \Gamma^{-1} = \gamma^{-1} \iota_s$. From the property that $K$ is an $s'$th order kernel, $Q_s(1)$ is a block diagonal matrix with first row $\iota_s$ and first column $\iota'_s$. $Q_s(1)^{-1}$ has the same property. Hence $\lim_{N \to \infty} \iota_s Q_s(y/\gamma)^{-1} = \iota_s$ and

$$\text{Var}[\iota_s \tilde{f}^{s-1}(y)] = \frac{1}{N} \gamma^{-1} f(y) \iota_s \int_{-1}^{1} \left[ K^{[s-1]}(w) K^{[s-1]}(w)^T \right] dw \iota'_s + o \left( \frac{1}{N \gamma} \right).$$

and the result follows. QED

**Remarks:**

1. Note that the asymptotic variance is the same as the usual formula when there is no boundary issue. It may be possible to get a more accurate measure of dispersion by using $Q_s(y/\gamma)$ in the calculations.

2. By inspection, the bias and variance vanish as $N \to \infty$, and so $\tilde{f}(y)$ is consistent in Mean Squared Error (MSE) and probability. Also by inspection, the rate of convergence in MSE is given by $\sqrt{\gamma^s + (N \gamma^{-1})^{-1}}$.

3. A biased reduced estimator of the $j$'th derivative of $f(y)$ is provided by $\tilde{f}^{(j)}(y) = \iota_{j+1} \tilde{f}^{[s-1]}(y)$ where $\iota_{j+1} = (0, \ldots, 0, 1, 0, \ldots, 0)$ and the one is the $j + 1$'th element of $\iota_{j+1}$. It follows that the bias of $\tilde{f}^{(j)}(y)$ is of order $O(\gamma^s)$. It is also straightforward to confirm that $\text{Var}[\tilde{f}^{(j)}(y)] = O((N \gamma^{1+2j})^{-1})$. The rate of convergence in MSE is given by $\sqrt{\gamma^s + (N \gamma^{1+2j})^{-1}}$.

4. Since each of the estimators are linear combinations of averages, it follows that the estimators, appropriately normalized (and with the appropriate conditions on the widow width) are asymptotically normal in distribution.

Consider a specific example using the Epanicheknikov kernel with $s = 1, 2$:

$$K(w) = \frac{3}{4}(1 - w^2)1[|w| < 1].$$

$$K^{[1]}(w) = \left( \frac{3}{4}(1 - w^2) \right) 1[|w| \leq 1]$$

$$Q_1(y/\gamma) = \int_{-1}^{y/\gamma} K^{(0)}(w) \, dw$$

$$= \int_{-1}^{y/\gamma} 3 \frac{1}{4}(1 - w^2) \, dw$$

$$= \frac{3}{4} \left[ \left( \frac{y}{\gamma} - \frac{1}{3} \left( \frac{y}{\gamma} \right)^3 \right) - \left( -1 - \frac{1}{3} (-1)^3 \right) \right]$$

$$= \frac{3}{4} \left[ \frac{2}{3} + \left( \frac{y}{\gamma} - \frac{1}{3} \left( \frac{y}{\gamma} \right)^3 \right) \right]$$

(2.11)

Note, for $y \geq \gamma$, $Q_1(1) = 1$. 

5
\[
Q_2(y/\gamma) = \int_{-1}^{y/\gamma} \left( \frac{K(w)}{K^{(1)}(w)} \right) (1 - w) \, dw \\
= \int_{-1}^{y/\gamma} \left( \frac{3}{4}(1 - w^2) - \frac{3}{4}(w - w^3) \right) \, dw \\
= \left( \frac{3}{4}(w - w^3) - \frac{3}{4}(w^2 - \frac{w^4}{4}) \right) \bigg|_{-1}^{y/\gamma} \\
= \left( \frac{3}{4} \left[ \frac{2}{3} + \left( \frac{y}{\gamma} - \frac{1}{3} \left( \frac{y}{\gamma} \right)^3 \right) \right] \right) \frac{3}{4} \left[ \frac{1}{4} - \left( \frac{1}{2} \left( \frac{y}{\gamma} \right)^2 - \frac{1}{4} \left( \frac{y}{\gamma} \right)^4 \right) \right] \right) \\
\] (2.12)

Note, for \( y \geq \gamma \), \( Q_2(1) \) is simply the 2 × 2 identity matrix.

### 3. Monte Carlo Study

Here we examine the performance of our bias reducing density estimation approach in the context of a small scale Monte Carlo experiment. We construct the data \( Y_i, i = 1, ..., N \), from an Exponential(1) distribution implying a left boundary of zero. We evaluate the performance of each density estimator over a mesh of 101 equally spaced points in the boundary region \([0, \gamma]\) where \( \gamma \equiv \gamma(N, K) \) is the smoothing parameter which is a function of both the sample size and the underlying kernel \( K \). We consider sample sizes \( N = 50, 100, 200, \) and 500. We consider two kernels: the quadratic kernel

\[
K_2(w) = \frac{3}{4} \left( 1 - w^2 \right) I_{(-1,1)}(w) \\
\] (3.1)

and the quartic kernel

\[
K_4(w) = \frac{15}{32} \left( 3 - 10w^2 + 7w^4 \right) I_{(-1,1)}(w) \\
\] (3.2)

where \( I_{(-1,1)}(w) \) is an indicator taking the value 1 if \( w \in (-1, 1) \) and zero otherwise. Each simulation is based on 500 replications.

For a given kernel function, \( K \), we denote our bias reducing density estimator with order of bias reduction \( s \) by \( \tilde{f}_s \). For the case of \( K_2 \) we consider \( s = 1, 2, 3 \) while for \( K_4 \) we consider \( s = 1, 2, 3, 4, 5 \).

For comparative purposes we also consider the typical fixed bandwidth density estimator

\[
\tilde{f}(y) = \frac{1}{N\gamma} \sum_{i=1}^{N} K \left( \frac{y - Y_i}{\gamma} \right) \\
\] (3.3)

and the adaptive density estimator

\[
f_A(y) = \frac{1}{N\gamma} \sum_{i=1}^{N} \frac{1}{\lambda_i} K \left( \frac{y - Y_i}{\gamma\lambda_i} \right) \\
\] (3.4)

\(^2\)See Gasser, Muller, and Mammitzsch (1985), p. 243, Table 1.
where $\lambda_i$ is a local bandwidth factor given by

$$\lambda_i = \left[ \frac{\widehat{f}(Y_i)}{\exp\left(\frac{1}{N} \sum_{i=1}^{N} \log \widehat{f}(Y_i)\right)} \right]^{0.5} \tag{3.5}$$

Since $\widehat{f}$ is not designed to perform well in finite samples with bounded data, we also consider an alternative that was designed for this case. In particular, we consider the boundary kernel approach of Gasser and Muller (1979). Let $K$ be a $k^{th}$ order polynomial kernel with support $[-1, 1]$. In our context where the data has a left boundary of zero, the Gasser-Muller boundary kernel can then be written as

$$f_{GM}(y) = \frac{1}{N\gamma} \sum_{i=1}^{N} K_q\left(\frac{y - Y_i}{\gamma}\right) \quad y \in [0, \gamma] \tag{3.6}$$

where

$$K_q(w) = \left(c_{0,q} + c_{1,q}w + \cdots + c_{k-1,q}w^{k-1}\right) K(w) I_{(-1,q)}(w) \tag{3.7}$$

is the “boundary kernel”, $q(y/\gamma) = \min\{1, y/\gamma\}$, and $c_{0,q}, c_{1,q}, \ldots, c_{k-1,q}$ are chosen to ensure that

$$\int_{-1}^{q(y/\gamma)} K_q(w) dw = 1$$
$$\int_{-1}^{q(y/\gamma)} w^j K_q(w) dw = \begin{cases} 0 & j = 1, 2, \ldots, k-1 \\ C < \infty & j = k \end{cases} \tag{3.8}$$

at each point $y$ in the boundary region $[0, \gamma]$ where the density is estimated. Thus the boundary kernel approach adjusts the kernel weights to ensure that the weighting function using in the boundary region satisfies the same moment restrictions as the $k^{th}$ order kernel. Note that when $y > \gamma$, $c_{0,q} = 1$ and $c_{1,q} = 0$, $c_{2,q} = 0$, ...$c_{k-1,q} = 0$ so that $f_{GM}(y)$ reduces to $\widehat{f}(y)$ for all points outside the boundary.

For each sample size and each optimal kernel, we choose the bandwidth $\gamma$ to minimize asymptotic mean integrated squared error of the fixed bandwidth kernel density estimator, $\widehat{f}$, under Exponential (1) data. While we could consider choosing $\gamma$ optimally for each density estimator, this would pose some problems for interpreting the results since the boundary region itself varies with $\gamma$.

The results for the third order bias reduction are mixed. Our proposed estimator dominates in that case in bias, but not overall in MSE. However, with the fifth order bias reduction our proposed estimator clearly dominates the others in terms of bias and mean squared error.

4. Conclusion

A method has been developed for boundary bias reduction of a variety of kernel estimators. These estimators are simple to compute, their asymptotic properties are comparable to the usual kernel estimators outside the boundary region and they performed well in the simulations we conducted.

---

References


### Table 1: Performance in the Boundary Region: Quadratic Kernel

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\gamma$</th>
<th>$\hat{f}$</th>
<th>$f_A$</th>
<th>$\tilde{f}_1$</th>
<th>$\tilde{f}_2$</th>
<th>$\tilde{f}_3$</th>
<th>$f_{GM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.9029</td>
<td>-0.0804</td>
<td>-0.0603</td>
<td>-0.0653</td>
<td>0.0102</td>
<td>0.0014</td>
<td>0.0183</td>
</tr>
<tr>
<td>100</td>
<td>0.7860</td>
<td>-0.0800</td>
<td>-0.0612</td>
<td>-0.0623</td>
<td>0.0084</td>
<td>0.0004</td>
<td>0.0138</td>
</tr>
<tr>
<td>200</td>
<td>0.6843</td>
<td>-0.0802</td>
<td>-0.0609</td>
<td>-0.0588</td>
<td>0.0065</td>
<td>-0.0008</td>
<td>0.0110</td>
</tr>
<tr>
<td>500</td>
<td>0.5697</td>
<td>-0.0786</td>
<td>-0.0587</td>
<td>-0.0514</td>
<td>0.0070</td>
<td>0.0015</td>
<td>0.0108</td>
</tr>
</tbody>
</table>

### Average Variance

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\gamma$</th>
<th>$\hat{f}$</th>
<th>$f_A$</th>
<th>$\tilde{f}_1$</th>
<th>$\tilde{f}_2$</th>
<th>$\tilde{f}_3$</th>
<th>$f_{GM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.9029</td>
<td>0.0086</td>
<td>0.0197</td>
<td>0.0039</td>
<td>0.0116</td>
<td>0.0217</td>
<td>0.0139</td>
</tr>
<tr>
<td>100</td>
<td>0.7860</td>
<td>0.0056</td>
<td>0.0119</td>
<td>0.0027</td>
<td>0.0076</td>
<td>0.0145</td>
<td>0.0090</td>
</tr>
<tr>
<td>200</td>
<td>0.6843</td>
<td>0.0035</td>
<td>0.0074</td>
<td>0.0018</td>
<td>0.0045</td>
<td>0.0085</td>
<td>0.0053</td>
</tr>
<tr>
<td>500</td>
<td>0.5697</td>
<td>0.0019</td>
<td>0.0040</td>
<td>0.0011</td>
<td>0.0026</td>
<td>0.0044</td>
<td>0.0029</td>
</tr>
</tbody>
</table>

### Average MSE

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\gamma$</th>
<th>$\hat{f}$</th>
<th>$f_A$</th>
<th>$\tilde{f}_1$</th>
<th>$\tilde{f}_2$</th>
<th>$\tilde{f}_3$</th>
<th>$f_{GM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>2.2680</td>
<td>-0.0686</td>
<td>-0.0538</td>
<td>-0.0140</td>
<td>0.0056</td>
<td>0.0062</td>
<td>-0.0063</td>
</tr>
<tr>
<td>100</td>
<td>2.0999</td>
<td>-0.0700</td>
<td>-0.0507</td>
<td>-0.0139</td>
<td>0.0050</td>
<td>0.0057</td>
<td>-0.0048</td>
</tr>
<tr>
<td>200</td>
<td>1.9442</td>
<td>-0.0710</td>
<td>-0.0472</td>
<td>-0.0132</td>
<td>0.0047</td>
<td>0.0051</td>
<td>-0.0030</td>
</tr>
<tr>
<td>500</td>
<td>1.7560</td>
<td>-0.0725</td>
<td>-0.0437</td>
<td>-0.0125</td>
<td>0.0045</td>
<td>0.0045</td>
<td>-0.0021</td>
</tr>
</tbody>
</table>

### Table 2: Performance in the Boundary Region: Quartic Kernel

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\gamma$</th>
<th>$\hat{f}$</th>
<th>$f_A$</th>
<th>$\tilde{f}_1$</th>
<th>$\tilde{f}_2$</th>
<th>$\tilde{f}_3$</th>
<th>$\tilde{f}_4$</th>
<th>$\tilde{f}_5$</th>
<th>$f_{GM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>2.2680</td>
<td>-0.0686</td>
<td>-0.0538</td>
<td>-0.0140</td>
<td>0.0056</td>
<td>0.0062</td>
<td>-0.0063</td>
<td>-0.0005</td>
<td>-0.0053</td>
</tr>
<tr>
<td>100</td>
<td>2.0999</td>
<td>-0.0700</td>
<td>-0.0507</td>
<td>-0.0139</td>
<td>0.0050</td>
<td>0.0057</td>
<td>-0.0048</td>
<td>-0.0003</td>
<td>-0.0047</td>
</tr>
<tr>
<td>200</td>
<td>1.9442</td>
<td>-0.0710</td>
<td>-0.0472</td>
<td>-0.0132</td>
<td>0.0047</td>
<td>0.0051</td>
<td>-0.0030</td>
<td>-0.0002</td>
<td>-0.0043</td>
</tr>
<tr>
<td>500</td>
<td>1.7560</td>
<td>-0.0725</td>
<td>-0.0437</td>
<td>-0.0125</td>
<td>0.0045</td>
<td>0.0045</td>
<td>-0.0021</td>
<td>-0.0002</td>
<td>-0.0024</td>
</tr>
</tbody>
</table>

### Average Variance

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\gamma$</th>
<th>$\hat{f}$</th>
<th>$f_A$</th>
<th>$\tilde{f}_1$</th>
<th>$\tilde{f}_2$</th>
<th>$\tilde{f}_3$</th>
<th>$\tilde{f}_4$</th>
<th>$\tilde{f}_5$</th>
<th>$f_{GM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>2.2680</td>
<td>0.0010</td>
<td>0.0020</td>
<td>0.0015</td>
<td>0.0027</td>
<td>0.0040</td>
<td>0.0552</td>
<td>0.0116</td>
<td>0.0488</td>
</tr>
<tr>
<td>100</td>
<td>2.0999</td>
<td>0.0006</td>
<td>0.0013</td>
<td>0.0009</td>
<td>0.0017</td>
<td>0.0024</td>
<td>0.0638</td>
<td>0.0071</td>
<td>0.0298</td>
</tr>
<tr>
<td>200</td>
<td>1.9442</td>
<td>0.0004</td>
<td>0.0008</td>
<td>0.0006</td>
<td>0.0010</td>
<td>0.0014</td>
<td>0.0192</td>
<td>0.0038</td>
<td>0.0168</td>
</tr>
<tr>
<td>500</td>
<td>1.7560</td>
<td>0.0002</td>
<td>0.0005</td>
<td>0.0003</td>
<td>0.0005</td>
<td>0.0007</td>
<td>0.0082</td>
<td>0.0018</td>
<td>0.0070</td>
</tr>
</tbody>
</table>

### Average MSE

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\gamma$</th>
<th>$\hat{f}$</th>
<th>$f_A$</th>
<th>$\tilde{f}_1$</th>
<th>$\tilde{f}_2$</th>
<th>$\tilde{f}_3$</th>
<th>$\tilde{f}_4$</th>
<th>$\tilde{f}_5$</th>
<th>$f_{GM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>2.2680</td>
<td>0.0437</td>
<td>0.0356</td>
<td>0.0110</td>
<td>0.0043</td>
<td>0.0044</td>
<td>0.0563</td>
<td>0.0116</td>
<td>0.0502</td>
</tr>
<tr>
<td>100</td>
<td>2.0999</td>
<td>0.0431</td>
<td>0.0351</td>
<td>0.0097</td>
<td>0.0029</td>
<td>0.0027</td>
<td>0.0372</td>
<td>0.0071</td>
<td>0.0305</td>
</tr>
<tr>
<td>200</td>
<td>1.9442</td>
<td>0.0426</td>
<td>0.0354</td>
<td>0.0085</td>
<td>0.0021</td>
<td>0.0016</td>
<td>0.0194</td>
<td>0.0038</td>
<td>0.0173</td>
</tr>
<tr>
<td>500</td>
<td>1.7560</td>
<td>0.0416</td>
<td>0.0365</td>
<td>0.0070</td>
<td>0.0012</td>
<td>0.0008</td>
<td>0.0083</td>
<td>0.0018</td>
<td>0.0072</td>
</tr>
</tbody>
</table>