Published by
NIGERIAN SOCIETY OF PHYSICAL SCIENCES
Available online @ https://journal.nsps.org.ng/index.php/jnsps

J. Nig. Soc. Phys. Sci. 6 (2024) 1631

Journal of the Nigerian Society of Physical Sciences

On the fourth-order hybrid beta polynomial kernels in kernel density estimation

Benson Ade Eniola Afere*

Department of Mathematical Sciences, Faculty of Natural Sciences, Prince Abubakar Audu University, 272102, Anyigba, Nigeria.

Abstract

This paper introduces a novel family of fourth-order hybrid beta polynomial kernels tailored for statistical analysis. The efficacy of these kernels is evaluated using two principal performance metrics: asymptotic mean integrated squared error (AMISE) and kernel efficiency. Comprehensive assessments were conducted using both simulated and real-world datasets, enabling a thorough comparison with conventional fourth-order polynomial kernels. The evaluation process entailed computing AMISE and efficiency metrics for both the hybrid and classical kernels. Consistently, the results illustrated the superior performance of the hybrid kernels over their classical counterparts across diverse datasets, underscoring the robustness and effectiveness of the hybrid approach. By leveraging these performance metrics and conducting evaluations on simulated and real-world data, this study furnishes compelling evidence supporting the superiority of the proposed hybrid beta polynomial kernels. The heightened performance, evidenced by lower AMISE values and elevated efficiency scores, strongly advocates for the adoption of the proposed kernels in statistical analysis tasks, presenting a marked improvement over traditional kernels.

DOI:10.46481/jnsps.2024.1631

Keywords: Kernel density estimation, Fourth-order kernels, hybrid kernels, AMISE, Efficiency.

Article History :
Received: 27 June 2023
Received in revised form: 15 December 2023
Accepted for publication: 24 January 2024
Published: 15 February 2024

Communicated by: T. Latunde

1. Introduction

In modern statistical practice, data visualization plays a crucial role in creating visual representations of data sets [1, 2]. One technique that has gained significant attention is kernel density estimation (KDE), which involves constructing a probability density estimate from observed data using parametric or nonparametric methods, or a combination of both (semi-parametric approach). Introduced by [3], and further development through the subsequent contributions of [4–6], KDE has become widely used in the statistical and scientific communities.

The parametric approach in KDE relies on assuming a known parametric family of distributions, requiring model building and prior knowledge of the data's underlying equations [7, 8]. However, this approach faces challenges when dealing with multimodal densities [9]. On the other hand, nonparametric methods, such as the nonparametric kernel density estimator (NKDE), allow the data to determine the density estimate without relying on specific assumptions. These methods are particularly useful when the population data has an unknown distribution or when the sample size is small. Nonparametric methods, including

*Corresponding author: Tel.: +234-913-539-8886
Email address: baafere3@gmail.com (Benson Ade Eniola Afere)

NKDE, are flexible and widely applicable, as they do not rely on rigid assumptions [10, 11].

NKDE is a robust and elegant smoothing method used in statistical and related problems [12, 13]. It finds applications in various fields, such as econometrics [14], insurance [15], climatology with emphasis on typhoon genesis [16], differential equations [17], and several others. KDE is a widely employed nonparametric technique used in statistics, data analysis, and machine learning to estimate the probability density function (PDF) of a random variable [6, 18]. The fundamental concept of kernel density estimation involves placing a kernel function at each data point and summing them up to approximate the PDF. The choice of the kernel function and bandwidth parameters significantly affects the quality of the density estimate [9, 18].

Commonly used kernel functions in KDE include the Gaussian, Epanechnikov, and biweight kernels, each with its own properties and impact on the estimated density [9, 18]. For instance, the Gaussian kernel is smooth and bell-shaped, while the Epanechnikov kernel has a flat central region and sharp tails [19].

Although selecting the optimal bandwidth poses challenges for the NKDE method, and there may be minor challenges associated with kernel functions, these factors do not diminish the importance of choosing and developing appropriate kernel functions [20]. In this context, this article proposes a family of fourth-order hybrid beta polynomial kernels. These kernels aim to address challenges related to higher AMISE in lower-order kernels and strive to improve the performance of kernel density estimation.

The subsequent sections of this article are organised as follows: Section 2 provides a comprehensive methodology, including an introduction to KDE and a discussion on error and efficiency criteria. This section covers both the second-order classical kernels and the proposed hybrid kernels. Section 3 focuses on the development of higher-order kernels. It outlines the construction process for both the fourth-order classical kernels and the newly proposed fourth-order hybrid kernels. In Section 4, we delve into the numerical verification of the results obtained. This section highlights the experimental procedures employed to evaluate the performance of the kernels, including Monte Carlo simulations and real-life experiments. Lastly, in Section 5, we present the discussion of the results and draw conclusions based on the findings obtained from the evaluations conducted in the previous sections. This section serves as a comprehensive summary of the article's main outcomes and their implications.

1.1. Preliminaries

To establish the foundation for this article, we introduce the following definition and lemma:

Definition 1.1. *For a given integer ℓ, let $k_\ell(t)$ be a user-chosen real-valued function defined for $t \in (-\infty, \infty)$ such that*

$$\int k_\ell(t)\,dt = 1. \tag{1}$$

We define the function $\mu_j(k_\ell(t))$ as follows:

$$\mu_j(k_\ell(t)) = \int t^j k_\ell(t)\,dt, \quad 0 \le j \le \ell. \tag{2}$$

The function $\mu_j(k_\ell(t))$ has the following properties:

$$\mu_j(k_\ell(t)) = \begin{cases} 1, & \text{if } j = 0, \\ 0, & \text{if } 1 \le j \le \ell - 1, \\ \mu_\ell \neq 0, & \text{if } j = \ell. \end{cases} \tag{3}$$

Remark 1. *The function $k_\ell(t)$ that satisfies Eq. (3) is referred to as an ℓ-th order symmetric kernel.*

Remark 2. *If $\ell = 2$, $k_2(t)$ is called a second-order kernel, and it is referred to as a lower-order kernel.*

Remark 3. *If $\ell \ge 4$, $k_\ell(t)$ is called a higher-order kernel, and it is referred to as a bias-reducing kernel. In particular, if $\ell = 4$, $k_4(t)$ is called a fourth-order kernel.*

Lemma 1.1. *Let $X_1, X_2, \cdots, X_n$ be a random sample of size p collected on a random variable X. If $t \in \mathbb{R}$ is a dummy variable, then as $p \to \infty$, $(1 - x^2)^p \to e^{-\frac{t^2}{2}}$.*

Proof. Consider a random sample $X_1, X_2, \cdots, X_n$ of size p from a random variable X with mean μ and variance σ^2. We define the random variable Z_p as follows:

$$Z_p = \frac{\frac{\bar{x}_p - \mu}{\sigma}}{\sqrt{p}}. \tag{4}$$

where,

$$\bar{x}_p = \frac{1}{p}\sum_{i=1}^{p} x_i. \tag{5}$$

As p tends to infinity, the distribution of Z_p converges to a standard normal distribution, $N(0, 1)$. Let's define the random variable U_i as:

$$U_i = \frac{x_i - \mu}{\sigma}. \tag{6}$$

Since the sample $X_1, X_2, \cdots, X_n$ is independent and identically distributed, the random variables U_i are also independent and identically distributed. Thus, we have $E(U_i) = 0$ and $Var(U_i) = 1$. The moment-generating function (mgf) of U_i can be expressed as:

$$m_{U_i}(t) = 1 + \frac{t^2}{2!}E(U_i^2) + \frac{t^3}{3!}E(U_i^3) + \cdots. \tag{7}$$

Now, considering Eqs. (4) and (6), we can write Z_p as:

$$Z_p = \frac{1}{\sqrt{p}}\sum i = 1^p U_i. \tag{8}$$

Since the X_i's are independent, the U_i's are also independent. Therefore, the mgf of U_i can be written as:

$$m_{U_i}(t) = \left(1 + \frac{t^2}{2!}E(U_i^2) + \frac{t^3}{3!}E(U_i^3) + \cdots\right)^p. \tag{9}$$

Taking the natural logarithm of both sides of Eq. (9), we get:

$$\ln m_{U_i}(t) = p \ln \left(\begin{array}{l} 1 + \frac{t^2}{2!} E(U_i^2) + \\ \frac{t^3}{3!} E(U_i^3) + \cdots \end{array} \right). \tag{10}$$

We can rewrite Eq. (10) as:

$$\ln m_{U_i}(t) = p \ln \left(\begin{array}{l} 1 - \left(-\frac{t^2}{2!} E(U_i^2) \right. \\ \left. \quad -\frac{t^3}{3!} E(U_i^3) - \cdots \right) \end{array} \right). \tag{11}$$

Now, let's consider $(1 - x^2)^p$. By taking the natural logarithm of this expression, we have:

$$\ln (1 - x^2)^p = p \ln(1 - x^2). \tag{12}$$

Expanding Eq. (12) using the Taylor series, we obtain:

$$\ln (1 - x^2)^p = p \left(-x^2 - \frac{x^4}{2} - \cdots \right). \tag{13}$$

Now, let $x^2 = -(\frac{t^2}{2p} + \frac{t^3}{3!p^{\frac{1}{2}}} E(U_i^3))$. Substituting this into Eq. (13), we get:

$$\ln m_{U_i}(t) = p \ln \left(1 - x^2\right).$$

This implies that:

$$\begin{aligned} \ln (1 - x^2)^p = p \left(-\left(\frac{t^2}{2!p} + \frac{t^3}{3!p^{\frac{3}{2}}} E(U_i^3) \right) \right) - \\ \frac{\left(-\left(\frac{t^2}{2!p} + \frac{t^3}{3!p^{\frac{3}{2}}} E(U_i^3) \right) \right)^2}{2} - \cdots . \end{aligned} \tag{14}$$

Simplifying Eq.(14), we have:

$$\begin{aligned} \ln (1 - x^2)^p = -\left(\frac{t^2}{2!} + \frac{t^3}{3!p^{\frac{1}{2}}} E(U_i^3) \right) - \\ \frac{\left(-\left(\frac{t^2}{2!p^{\frac{1}{2}}} + \frac{t^3}{3!p} E(U_i^3) \right) \right)^2}{2} - \cdots . \end{aligned}$$

Thus, as $p \to \infty$, $\ln (1 - x^2)^p \to -\frac{t^2}{2}$, and therefore, $(1 - x^2)^p \to e^{-\frac{t^2}{2}}$ which completes the proof. □

2. Methodology

2.1. Kernel Density Estimation

The KDE is a nonparametric method used to estimate the PDF of a random variable based on a sample of observations. Given a random sample $X_1, X_2, \cdots, X_n$ of size n collected from a random variable X with probability distribution function $F(x)$, the KDE estimates the PDF $f(x)$ at a point x using the following formula:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} k\left(\frac{x - X_i}{h} \right). \tag{15}$$

where, h is the smoothing parameter, commonly referred to as the bandwidth. The bandwidth controls the width of the kernel function and influences the smoothness of the estimated density. A smaller bandwidth leads to a rougher estimate with more variability, while a larger bandwidth produces a smoother estimate with reduced variability.

The kernel function $k(t)$ determines the shape of the density estimation at each data point, and it must satisfy certain properties. Typically, the kernel function is chosen to be a symmetric unimodal density function which is the desired properties of the estimated density. Commonly used kernel functions include the Gaussian, Epanechnikov, and biweight kernels [9, 18]. These kernel functions must satisfy the following axioms:

$$\left. \begin{array}{ll} \text{i.} & \int k(t)dt = 1, \\ \text{ii.} & \int tk(t)dt = 0, \\ \text{iii.} & \int t^2 k(t)dt \neq 0. \end{array} \right\} \tag{16}$$

These axioms guarantee that the kernel function integrates to unity, has zero mean, and a non-zero second moment. These properties enable accurate smoothing and preserve the essential characteristics of the estimated density. By adjusting the bandwidth parameter h and selecting an appropriate kernel function $k(t)$, the kernel density estimator provides a flexible and versatile tool for estimating probability density functions from empirical data.

2.2. Error Criterion

As mentioned earlier, the kernel function is typically chosen to be a non-negative density function that is symmetric around zero. However, recent research has shown that relaxing the conditions for the kernel function can lead to better performance. In this paper, we consider kernel functions that satisfy the following axioms, as discussed in Refs. [9, 20, 21] (building on the arguments in Refs. [6, 18]):

$$\begin{array}{ll} \text{i.} & \int k(t)dt = 1, \\ \text{ii.} & \int tk(t)dt = \int t^2 k(t)dt = \int t^3 k(t)dt = 0, \\ \text{iii.} & \int t^4 k(t)dt \neq 0. \end{array} \tag{17}$$

Previous studies have mainly focused on reducing global errors, as demonstrated in works [8, 22]. Thus, the primary objective of this paper is to introduce a new set of kernel functions that can significantly reduce global errors. We propose the use of fourth-order hybrid beta polynomial kernels, which form a family of kernel functions specifically designed to minimize global errors. This shall be constructed in Subsection 3.1.2.

Theorem 2.1. *Let f be a sufficiently smooth and bounded density function, and let $k(t)$ satisfy the conditions in Eq. (17). If x is a point with $f(x) > 0$, and f is continuously differentiable up to the fourth order in a neighbourhood of x, then the asymptotic mean integrated squared error (AMISE) of any fourth-order kernel can be expressed as:*

$$AMISE\hat{f}_h(x) = \frac{h^8}{(4!)^2}\int (f^{(4)}(x))^2 dx\left(\int t^4 k(t)dt\right)^2 + \frac{1}{nh}\int k^2(t)dt.$$

Proof. See Ref. [8] □

The foundation for our performance metrics lies in the theoretical framework established by Theorem 2.1. Building upon this, we further develop a comprehensive and generalized AMISE, as detailed and substantiated in the subsequent theoretical exposition provided in Theorem 2.2.

Theorem 2.2. *Expanding on Theorem 2.1, if we define k_{2m+2} as the $(2m+2)^{th}$-moment of any fourth-order kernel, and if f is continuously differentiable up to the $(2m+2)^{th}$-order in a neighborhood of x, then the generalized AMISE of any fourth-order kernel can be formulated as follows:*

$$AMISEE^{2m} = \left(\frac{4m+5}{4m+4}\right)\left(\frac{4m+4}{((2m+2)!)^2}\right) \int (f^{(2m+2)})^2 dx^{\frac{1}{4m+5}}((k_{2m+2})^2)^{\frac{1}{4m+5}}\left(\int k^2(t)dt\right)^{\frac{4m+4}{4m+5}} n^{-\frac{4m+4}{4m+5}}.$$

Proof. The global error for evaluating the discrepancy between the estimated density and the true density, as pointed out by [8], is given by:

$$\mathrm{MISE}\hat{f}_h(x) = \int \mathrm{bias}^2\hat{f}_h(x)dx + \int \mathrm{var}\hat{f}_h(x)dx. \tag{18}$$

where

$$\mathrm{bias}\hat{f}_h(x) = \mathbb{E}\hat{f}_h(x) - f(x). \tag{19}$$

and

$$\mathrm{var}\hat{f}_h(x) = \mathbb{E}^2\hat{f}_h(x) - (\mathbb{E}\hat{f}_h(x))^2. \tag{20}$$

Assuming that the underlying density is smooth enough and the kernel's fourth moment is finite, it is possible to use Taylor's series expansion to obtain the global error scheme of the estimator, which can be expressed as follows: Using Eq. (19), we have:

$$\mathrm{bias}\hat{f}_h(x) = \int \hat{f}_h(x)f(y)dy - f(x). \tag{21}$$

Substituting Eq. (15) into Eq. (21), we have:

$$\mathrm{bias}\hat{f}_h(x) = \frac{1}{h}\int k\left(\frac{x-y}{h}\right)f(y)dy - f(x). \tag{22}$$

Let $t = (x-y)/h$, then Eq. (22) becomes:

$$\mathrm{bias}\hat{f}_h(x) = \int k(t)\, f(x-th)dt - f(x). \tag{23}$$

Using Taylor series expansion up to the $(2m+2)^{th}$ - order on Eq. (23), we have:

$$\mathrm{bias}\hat{f}_h(x) = \int k(t)[f(x) - thf^{(1)}(x) + \frac{1}{2!}t^2h^2f^{(2)}(x) - \frac{1}{3!}t^3h^3f^{(3)}(x) + \cdots - \frac{1}{(2m+1)!}t^{2m+1}h^{2m+1}f^{(2m+1)}(x) + \frac{1}{(2m+2)!}t^{2m+2}h^{2m+2}f^{(2m+2)}(x) + \cdots - \cdots]dt - f(x).$$

Simplifying this, we have:

$$\mathrm{bias}\hat{f}_h(x) = (f(x)\int k(t)\,dt - hf^{(1)}(x)\int tk(t)dt + \frac{1}{2!}h^2f^{(2)}(x)\int t^2k(t)dt - \frac{1}{3!}h^3f^{(3)}(x)\int t^3k(t)dt + \cdots - \frac{1}{(2m+1)!}h^{2m+1}f^{(2m+1)}(x)\int t^{2m+1}k(t)dt + \frac{1}{(2m+2)!}\times h^{2m+2}f^{(2m+2)}(x)\int t^{2m+2}k(t)dt + \cdots) - f(x). \tag{24}$$

Applying the conditions in Eq. (17) and the $(2m+2)^{th}$ - moment (k_{2m+2}) in the statement of the theorem to Eq. (24), we have:

$$\mathrm{bias}\hat{f}_h(x) = \frac{1}{(2m+2)!}h^{2m+2}f^{(2m+2)}(x)(k_{2m+2})^2 + o(h^{2m+2}). \tag{25}$$

Similarly, we can derive the expression for the variance term. From Eq. (20), $\mathrm{var}\hat{f}_h(x)$ is given as:

$$\mathrm{var}\hat{f}_h(x) = \frac{1}{nh^2}\int k^2\left(\frac{x-y}{h}\right)f(y)dy. \tag{26}$$

As in the case bias, let $t = (x-y)/h$, hence Eq. (26) becomes:

$$\mathrm{var}\hat{f}_h(x) = \frac{1}{nh^2}\int k^2(t)f(x-th)dt. \tag{27}$$

Thus, using the same Taylor's series expansion as in bias term, Eq. (27) becomes:

$$\mathrm{var}\hat{f}_h(x) = \frac{1}{nh}\int k^2(t)[f(x) - thf^{(1)}(x) + o(h)]dt = \frac{1}{nh}f(x)\int k^2(t)dt + o\left(\frac{1}{nh}\right).$$

Hence, the $\mathrm{var}\hat{f}_h(x)$ is given as:

$$\mathrm{var}\hat{f}_h(x) = \frac{1}{nh}f(x)\int k^2(t)dt + o\left(\frac{1}{nh}\right). \tag{28}$$

Substituting Eqs.(25) and (28) into Eq. (18), we have:

$$\mathrm{MISE}\hat{f}_h(x) = \frac{h^{4m+4}}{((2m+2)!)^{2m+2}}\times \int ((f^{(2m+2)}(x))^2 dx)(k_{2m+2})^2 + \frac{1}{nh}\times \int k^2(t)dt + o(h^{4m+4}) + o\left(\frac{1}{nh}\right). \tag{29}$$

Simplifying further, we neglect the higher-order terms in Eq. (29) and obtain the AMISE as follows:

$$\text{AMISE}\hat{f}_h(x) = \frac{h^{4m+4}}{((2m+2)!)^{2m+2}} \times \int ((f^{(2m+2)}(x))^2 dx)(k_{2m+2})^2 + \frac{1}{nh}\int k^2(t)dt. \tag{30}$$

We differentiate Eq. (30) with respect to h and solving for h by setting the resultant differential equation to zero, we have:

$$h^{2m}_{\text{AMISE}} = \left(\frac{((2m+2)!)^2}{4m+4} \cdot \int (f^{(2m+2)}(x))^2 dx\right)^{\frac{1}{4m+5}} \times ((k_{2m+2})^2)^{\frac{1}{4m+5}} \left(\int k^2(t)dt\right)^{\frac{1}{4m+5}} n^{-\frac{1}{4m+5}}. \tag{31}$$

Now, put Eq. (31) into Eq. (30) and simplify completely, then we have the generalized AMISE that is free of h-the optimal bandwidth. And thus the generalized AMISE is given by:

$$\text{AMISE}^{2m} = \left(\frac{4m+5}{4m+4}\right)\left(\frac{4m+4}{((2m+2)!)^2}\int (f^{(2m+2)})^2 dx\right)^{\frac{1}{4m+5}} \times ((k_{2m+2})^2)^{\frac{1}{4m+5}} \left(\int k^2(t)dt\right)^{\frac{4m+4}{4m+5}} n^{-\frac{4m+4}{4m+5}}.$$

This completes the proof. □

2.3. Efficiency Criterion

Past research has predominantly focused on investigating the efficiency of univariate kernels [8, 23]. While [24] have explored the efficiency of Gaussian and biweight kernels, their emphasis was primarily on second-order univariate kernels.

In this article, our objective is twofold. First, we aim to calculate the global error of the fourth-order beta polynomial kernel. Second, we seek to examine its efficiency by comparing it with the Epanechnikov kernel, known for yielding the minimum asymptotic mean integrated squared error (AMISE) [18].

Efficiency in the context of symmetric kernels is denoted as $Eff(K_{sk})$ and defined by Eq. (32). It represents the ratio of the AMISE of the Epanechnikov kernel to that of any other symmetric kernel.

$$\text{Eff}(K_{sk}) = \left(\frac{C(K_e)}{C(K_{sk})}\right)^{\frac{5}{4}}. \tag{32}$$

Here, $C(K_{sk})$ is a constant associated with the kernel under consideration, determined by $(\int t^2k(t)dt)^{\frac{2}{5}}(\int K^2(t)dt)^{\frac{4}{5}}$. $C(K_e)$ represents the Epanechnikov kernel constant. The general expression for the efficiency of such second-order univariate kernels is defined by:

$$\text{Eff}^{2m}(K_{sk}) = \left(\frac{C_{2m}(K_e)}{C_{2m}(K_{sk})}\right)^{\frac{4m+1}{4m}}. \tag{33}$$

where $C_{2m}(K_{sk}) = (\int t^{2m}k(t)dt)^{\frac{2}{4m+1}}(\int k^2(t)dt)^{\frac{4m}{4m+1}}$ is the generalized higher-order constant of any second-order beta polynomial kernel. The expression $\int k^2(t)dt$ corresponds to the L_2 norm, whereas $\int t^{2m}k(t)dt$ signifies the $2m^{th}$ moments of second-order symmetric beta polynomial kernels, as defined in Eqs. (36) and (48). $C_{2m}(K_e)$ in Eq. (33) denotes the generalized higher-order constant of the second-order Epanechnikov kernel. However, if we employ the regularity conditions in Eq. (17), the efficiency equation becomes:

$$\text{Eff}(K_{sk}) = \left(\frac{C(K_e)}{C(K_{sk})}\right)^{\frac{9}{8}}. \tag{34}$$

Thus, the generalized efficiency scheme for evaluating the efficiency of any fourth-order beta polynomial kernel, by extending the efficiency equation in Eq. (34), is given by:

$$\text{Eff}^{2m+2}(K_{sk}) = \left(\frac{C_{2m+2}(K_e)}{C_{2m+2}(K_{sk})}\right)^{\frac{4m+5}{4m+4}}. \tag{35}$$

where $C_{2m+2}(K_{sk}) = (\int t^{2m+2}k(t)dt)^{\frac{2}{4m+5}}(\int k^2(t)dt)^{\frac{4m+4}{4m+5}}$ is the generalized higher-order constant of any fourth-order beta polynomial kernel. The L_2-norm and $(2m+2)^{th}$ moments are as defined previously.

2.4. A Family of Second-Order Classical Beta Polynomial Kernels

The family of second-order classical beta polynomial kernels is an important family of kernel functions used in kernel density estimation for data visualization [20]. These kernels provide smoother estimates with more derivatives as the degree of the function increases, and they have been widely used in various applications of KDE [17, 25, 26]. They offer flexibility in shaping the density estimate and can capture different patterns and features in the data. The general form of second-order classical beta polynomial kernels is given by the formula:

$$k_{[2,p]}(t) = \begin{cases} b_1(1-t^2)^p, & |t| \le 1, \\ 0, & \text{elsewhere.} \end{cases} \tag{36}$$

where $b_1 = \frac{(2p+1)!!}{2^{p+1}p!}$, p is the power of the family and takes values $p = 0, 1, 2, \cdots$. The range of integration is $[-1, 1]$, and the double factorial $(2p+1)!! = (2p+1)(2p-1)\cdot 5 \cdot 3 \cdot 1$. Different values of p produce different kernel functions. For example, $p = 0$ produces a uniform kernel, $p = 1$ produces a second-order Epanechnikov kernel, $p = 2$ produces a biweight kernel, $p = 3$ produces a triweight kernel, $p = 4$ produces a quadriweight kernel, and so on. These kernel functions are given in Eqs.(37), (38), (39), (40), and (41), respectively. As $p \to \infty$, Eq.(36) produces a second-order Gaussian kernel, given by Eq.(42).

$$k_{[2,0]}(t) = \frac{1}{2}(1-t^2)^0. \tag{37}$$

$$k_{[2,1]}(t) = \frac{3}{4}(1-t^2)^1. \tag{38}$$

$$k_{[2,2]}(t) = \frac{15}{16}(1-t^2)^2. \tag{39}$$

$$k_{[2,3]}(t) = \frac{35}{32}(1-t^2)^3. \tag{40}$$

$$k_{[2,4]}(t) = \frac{315}{256}(1-t^2)^4. \tag{41}$$

$$k_{[2,\phi]}(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}. \tag{42}$$

Eq. (36) sets the foundation for the construction of fourth-order classical beta polynomial kernels in Section 3.

2.5. A Family of Second-Order Hybrid Beta Polynomial Kernels

In a previous work by Afere [27], a family of second-order hybrid beta polynomial kernels was introduced using the following formula:

$$\begin{aligned} k_{[2,h]}(t) &= \rho_1 k_{[2,p]}(t) + \\ &\quad \rho_2 k_{[2,p+1]}(t); \quad 0 < \rho_1 < 1, \quad \rho_1 + \rho_2 = 1. \end{aligned} \tag{43}$$

Here, $k_{[2,p]}(t)$ is defined as in Eq. (36), and $k_{[2,p+1]}(t)$ is the $(p+1)^{st}$ power of $k_{[2,p]}(t)$. When we set $\rho_1 = \frac{1}{2}$, Eq. (43) simplifies to:

$$k_{[2,h]}(t) = \frac{1}{2}k_{[2,p]}(t) + \frac{1}{2}k_{[2,p+1]}(t). \tag{44}$$

The $(p+1)^{st}$ power of $k_{[2,p]}(t)$ is defined as:

$$k_{[2,p+1]}(t) = \begin{cases} b_2(1-t^2)^{p+1}, & |t| \le 1, \\ 0, & \text{elsewhere.} \end{cases} \tag{45}$$

where $b_2 = \frac{(2p+3)!!}{2^{p+2}(p+1)!}$, $(2p+3)!! = (2p+3)(2p+1)\cdot\ldots\cdot 3\cdot 1$ denotes the double factorial. Substituting Eqs.(36) and (45) into Eq. (44) and simplifying, we obtain:

$$k_{[2,h]}(t) = a_1(t,p)\frac{b_1}{4(p+1)}(1-t^2)^p. \tag{46}$$

where $a_1(t,p) = [(4p+5) - (2p+3)t^2]$. By letting $k_{[2,h]}(t) = k^h_{[2,p]}(t)$, Eq. (46) becomes:

$$k^h_{[2,p]}(t) = \begin{cases} \frac{a_1(t,p)b_1}{4(p+1)}(1-t^2)^p, & |t| \le 1, \\ 0, & \text{elsewhere.} \end{cases} \tag{47}$$

The family of kernel functions in Eq.(47) can be further generalized using arithmetic progression, leading to:

$$k^h_{[2,p]}(t) = \begin{cases} b_1 a_2\left(1 - a_3 t^2\right)\times \\ \qquad (1-t^2)^p, & |t| \le 1, \\ 0, & \text{elsewhere.} \end{cases} \tag{48}$$

where $a_2 = \frac{4+\frac{5}{p}}{4+\frac{4}{p}}$ and $a_3 = \frac{2+\frac{3}{p}}{4+\frac{4}{p}}$. The specific kernel functions for $p=1$, $p=2$, $p=3$, and $p=4$ can be written as follows:

$$\begin{aligned} k^h_{[2,1]}(t) &= \frac{1}{8}(9-5t^2)k_{[2,1]}(t), \\ k^h_{[2,2]}(t) &= \frac{1}{12}(13-7t^2)k_{[2,2]}(t), \\ k^h_{[2,3]}(t) &= \frac{1}{16}(17-9t^2)k_{[2,3]}(t), \\ k^h_{[2,4]}(t) &= \frac{1}{20}(21-11t^2)k_{[2,4]}(t). \end{aligned} \tag{49}$$

As $p \to \infty$, we have $b_1 \to \frac{2}{\sqrt{2\pi}}$, $a_2 \to 1$, $a_3 \to \frac{1}{2}$, and by Lemma 1.1, $(1-t^2)^p \to e^{-\frac{1}{2}t^2}$. Therefore, Eq. (48) tends to:

$$k^h_{[2,\phi]}(t) = \begin{cases} \left(2-t^2\right)k_{[2,\phi]}(t), & t \in (-\infty,\infty), \\ 0, & \text{elsewhere.} \end{cases} \tag{50}$$

Eq. (45) represents the family of second-order hybrid beta polynomial kernels, which will be utilized in Section 3.1.2 to construct the proposed family of fourth-order hybrid beta polynomial kernels. When we combine Eqs.(48) and (50), we obtain the following expression for $k^h_{[2,p]}(t)$ as:

$$k^h_{[2,p]}(t) = \begin{cases} b_1 a_2(1-a_3t^2)(1-t^2)^p, & p < \infty, \\ (2-t^2)k_{[2,\phi]}(t), & p \to \infty. \end{cases} \tag{51}$$

On using the properties of second-order kernels in Eq. (16) on Eq. (51), we have:

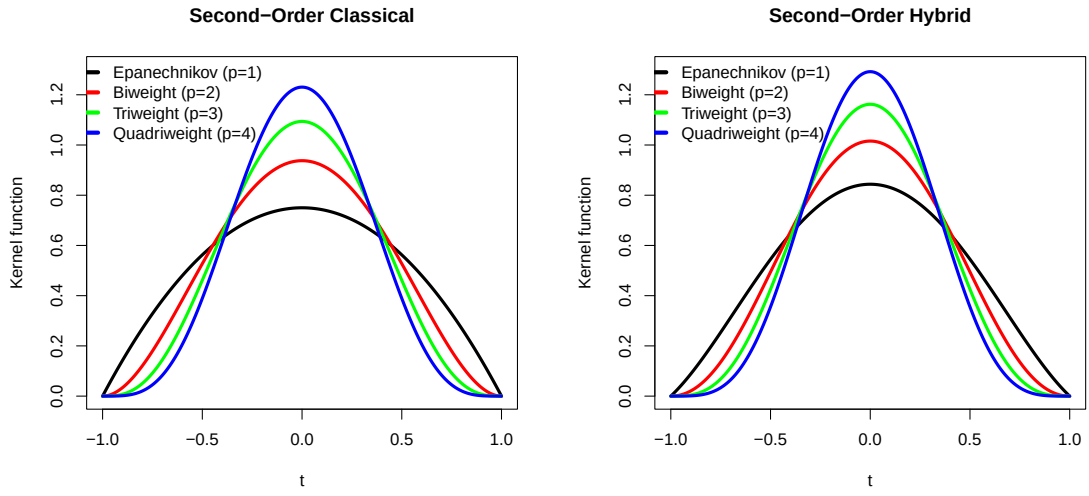
$$\int k(t)dt = \begin{cases} \frac{b_1\sqrt{\pi}\,\Gamma(p+2)}{(p+1)\Gamma(p+\frac{3}{2})} = 1, & \text{if } p < \infty, \\ 1, & \text{if } p \to \infty. \end{cases} \tag{52}$$

$$\int tk(t)dt = \begin{cases} 0, & \text{if } p < \infty, \\ 0, & \text{if } p \to \infty. \end{cases} \tag{53}$$

$$\int t^2k(t)dt = \begin{cases} \frac{b_1\sqrt{\pi}\Gamma(p+3)}{2(p+1)\Gamma(p+\frac{7}{2})} \neq 0, & \text{if } p < \infty, \\ -1 \neq 0, & \text{if } p \to \infty. \end{cases} \tag{54}$$

Eqs. (52)-(54) demonstrate that the family of second-order hybrid kernels exhibits the following characteristics: they integrate to unity, possess a zero mean, and have a non-zero second moment. This observation underscores that every kernel within this family possesses symmetry, as it satisfies all the criteria outlined in Eq. (16). To provide a visual representation of the kernels, Figure 1(a) and (b) present the graphs of the second-order classical beta polynomial kernels and the second-order hybrid beta polynomial kernels for $p=1$, $p=2$, $p=3$, and $p=4$. Figure 1 consists of two panels that showcase the shapes of the second-order classical beta polynomial kernels and the proposed second-order hybrid beta polynomial kernels. In panel (a), (Figure 1(a)), the curves represent the second-order classical beta polynomial kernels, where each curve corresponds to a specific value of the parameter p. Analyzing these curves allows us to observe the variations in the shape of the classical kernels as p changes. Thus, as p varies, the shape of the kernel changes. The curves exhibit a characteristic bell-shaped profile, which is a common feature of many kernel functions. With increasing values of p, the kernels become narrower and taller. This indicates that as p increases, the kernel concentrates its mass around the central point, making it more peaked. This visualization helps in comprehending the influence of different p values on the kernel shape.

In Figure 1(b), we observe the second-order hybrid beta polynomial kernels, each curve representing a specific value of p. These hybrid kernels, akin to their classical counterparts, exhibit varying shapes in response to changes in p. Notably, for



(a) (b)

Figure 1: Plot of shapes of: (a) second-order classical beta polynomial kernels and (b) proposed second-order hybrid beta polynomial kernels for different values of p.

equivalent p values, the proposed hybrid kernels appear wider and shorter compared to the classical ones. This indicates that the hybrid approach tends to yield kernels with a broader spread and flatter profile.

Comparing the curves in Figure 1(b) to those in Figure 1(a) allows us to discern the distinct shapes introduced by the hybridization process. The discernible differences in shape emphasize the influence of the hybridization on the behavior of the kernels.

Overall, Figure 1 provides a concise visual comparison between the shapes of second-order classical and hybrid beta polynomial kernels. This comparison aids in understanding how the hybridization process affects the kernel shape, which, in turn, assists in assessing their suitability for various applications, including density estimation.

3. Higher-Order Kernels

In Simonoff's investigation [13], the issue of bias contribution to the mean integrated squared error was addressed by introducing the concept of selecting a kernel function. Traditionally, kernel functions are required to have a second moment greater than zero [8]. However, by relaxing this requirement and setting the second moment of the kernel function to zero, it becomes possible to construct a kernel function with a non-zero fourth moment. This modification helps to reduce the bias to the order of h^4.

It is important to note that this approach is applicable under certain conditions for the density function f. Firstly, f should have a continuous distribution [5]. Additionally, f must possess a squared integrable fourth moment, meaning that the integral of f^2 over its entire range should converge [13]. Finally, the density function f should be monotonic, indicating that it consistently increases or decreases without fluctuations [28]. By satisfying these conditions and adopting the modified kernel function, the bias contribution to the mean integrated squared error can be reduced significantly [13].

3.1. Construction of Higher-Order Kernels

To construct higher-order kernels, we can employ the concept of bias-reducing kernels, which combine the characteristics of both negative and nonnegative kernels and can lead to faster convergence rates [21, 29]. Several rules for generating such kernels have been proposed in the literature, including those presented by [30]. In this paper, we will revisit the rule presented by [31], which involves starting with an ℓ-th order Gaussian kernel $k_{[\ell]}(t)$ defined as:

$$k_{[\ell]}(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}. \tag{55}$$

The first and second derivatives of $k_{[\ell]}(t)$ are given by:

$$k'_{[\ell]}(t) = -t\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}. \tag{56}$$

$$k''_{[\ell]}(t) = \frac{1}{\sqrt{2\pi}}(t^2 - 1)e^{-\frac{1}{2}t^2}. \tag{57}$$

By evaluating these derivatives at $t = 0$, we obtain:

$$k_{[\ell]}(0) = \frac{1}{\sqrt{2\pi}}. \tag{58}$$

$$k'_{[\ell]}(0) = 0. \tag{59}$$

$$k''_{[\ell]}(0) = -\frac{1}{\sqrt{2\pi}}. \tag{60}$$

Expanding $k_{[\ell]}(t)$ using the univariate Taylor series expansion up to order 2, we have:

$$\hat{k}_{[\ell]}(t) = \frac{1}{0!}k_{[\ell]}(0) + \frac{1}{1!}k'_{[\ell]}(0) + \frac{1}{2!}k''_{[\ell]}(0). \tag{61}$$

Substituting Eqs. (58), (59), and (60) into Eq. (61), we have:

$$\begin{aligned}\hat{k}_{[\ell]}(t) &= k_{[\ell]}(0) - \frac{1}{2}k''_{[\ell]}(0),\\ &= \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2} - \frac{1}{2}(t^2-1)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2},\\ &= \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2} - \frac{1}{2}t^2\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2} + \frac{1}{2}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2},\\ &= \left(1+\frac{1}{2}\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2} - \frac{1}{2}t^2\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2},\\ &= \frac{3}{2}k_{[\ell]}(t) + \frac{1}{2}t(-tk_{[\ell]}(t)),\\ &= \frac{3}{2}k_{[\ell]}(t) + \frac{1}{2}tk'_{[\ell]}(t), \text{ where } k'_{[\ell]}(t) = -tk_{[\ell]}(t).\end{aligned}$$

Thus, using Definition 1.1, we obtain the generalized formula:

$$k_{[\ell+2,p]}(t) = \frac{3}{2}k_{[\ell,p]}(t) + \frac{1}{2}tk'_{[\ell,p]}(t). \tag{62}$$

Here, $k_{[\ell+2,p]}(t) = \hat{k}_{[\ell]}(t)$. Eq. (62) provides a generalized formula for constructing $(\ell+2)^{nd}$ order kernels with p as the power of the kernel family. We will utilize this formula in Subsections 3.1.1 and 3.1.2 to construct the family of fourth-order classical polynomial kernels and the proposed family of fourth-order hybrid polynomial kernels, respectively.

3.1.1. *Construction of Fourth-Order Classical Beta Polynomial Kernels*

In this subsection, we will construct the family of fourth-order classical beta polynomial kernels using the $(p+1)^{\text{st}}$ power of Eq. (36). By taking $\ell = 2$, Eq. (62) can be rewritten as:

$$k_{[4,p]}(t) = \frac{3}{2}k_{[2,p+1]}(t) + \frac{1}{2}tk'_{[2,p+1]}(t). \tag{63}$$

To find $k'_{[2,p+1]}(t)$, we differentiate the $(p+1)^{st}$ power of Eq. (36) with respect to t and obtain:

$$k'_{[2,p+1]}(t) = -t\frac{(2p+3)!}{2^{p+1}(p)!}(1-t^2)^p. \tag{64}$$

Substituting Eqs. (36) and (64) into Eq. (63), we have:

$$k_{[4,p]}(t) = \frac{3}{2}\left(b_2(1-t^2)^{p+1}\right) + \frac{1}{2}t\left(-\frac{b_2(1-t^2)^p}{p+1}\right).$$

Simplifying this expression, we obtain:

$$k_{[4,p]}(t) = \begin{cases}\frac{b_2}{2}a_4(t,p)(1-t^2)^p, & \text{if } |t| \le 1,\\ 0, & \text{elsewhere.}\end{cases} \tag{65}$$

where $a_4(t,p) = (3-(2p+5)t^2)$. Furthermore, by simplifying Eq. (65) and using arithmetic progression, the family of kernel functions in Eq. (65) can be generalized as:

$$k_{[4,p]}(t) = \begin{cases}2p(p+1)b_1\times & \\ a_5(3-a_6t^2)(1-t^2)^p, & |t| \le 1,\\ 0, & \text{elsewhere.}\end{cases} \tag{66}$$

where $a_5 = \frac{2+\frac{3}{p}}{4+\frac{4}{p}}$ and $a_6 = \frac{2+\frac{5}{p}}{2}$. Eq. (66) represents the family of fourth-order classical beta polynomial kernels. Setting $\ell = 2$ and using the $(p+1)^{\text{st}}$ power of Eq. (36) leads to this expression. When p takes on the values 1, 2, 3, and 4, Eq. (66) gives rise to the fourth-order Epanechnikov, biweight, triweight, and quadriweight kernels, respectively. These kernels are expressed as follows:

$$k_{[4,1]}(t) = \frac{5}{8}(3-7t^2)k_{[2,1]}(t). \tag{67}$$

$$k_{[4,2]}(t) = \frac{7}{12}(3-9t^2)k_{[2,2]}(t). \tag{68}$$

$$k_{[4,3]}(t) = \frac{9}{16}(3-11t^2)k_{[2,3]}(t). \tag{69}$$

$$k_{[4,4]}(t) = \frac{11}{20}(3-13t^2)k_{[2,4]}(t). \tag{70}$$

As $p \to \infty$, $\frac{(2p+1)!!}{2^p(p-1)!} \to \frac{1}{\sqrt{2\pi}}$, $a_5 \to \frac{1}{2}$, $a_6 \to 1$, and by Lemma 1.1, $(1-t^2)^p \to e^{-\frac{1}{2}t^2}$. Therefore, Eq. (63) tends to:

$$k_{[4,\phi]}(t) = \begin{cases}\frac{1}{2}\left(3-t^2\right)k_{[2,\phi]}(t), & \text{if } t \in (-\infty,\infty),\\ 0, & \text{elsewhere.}\end{cases} \tag{71}$$

For further details, see [32]. The graph of Eqs.(67), (68), (69), and (70) is shown in Figure2(a).

3.1.2. *Construction of the Proposed Family of Fourth-Order Hybrid Beta Polynomial Kernels*

In this subsection, we will construct the proposed family of fourth-order hybrid beta polynomial kernels by using the kernels previously constructed in Subsection 2.5. We begin by differentiating the family of second-order hybrid beta polynomial kernels in Eq. (45) to obtain:

$$\left(k^h_{[2,p]}\right)'(t) = -t\frac{[(4p+3)-(2p+3)t^2](2p+1)!!}{2^{p+2}(p)!}(1-t^2)^{p-1}. \tag{72}$$

Now, substituting Eqs.(48) and (72) into Eq. (63), we have:

$$k^h_{[4,p]}(t) = \frac{3}{2}\left(\frac{a_1(t,p)(2p+1)!!}{2^{p+3}(p+1)!}(1-t^2)^p\right) + \frac{1}{2}t\left(-t\frac{[(4p+3)-(2p+3)t^2](2p+1)!!}{2^{p+2}(p)!}(1-t^2)^{p-1}\right). \tag{73}$$

Simplifying this, the proposed family of fourth-order hybrid beta polynomial kernels is given as:

$$k^h_{[4,p]}(t) = \begin{cases}\frac{b_1}{8(p+1)}(\alpha(t,p)- & \\ \beta(t,p))(1-t^2)^{p-1}, & |t| \le 1,\\ 0, & \text{elsewhere.}\end{cases} \tag{74}$$

where b_1 is as given in Eq. (36), $\alpha(t,p) = 2(1+2p)!(1+p)(3-(3+2p)t^2)$ and $\beta(t,p) = (3+2p)(3-(5+2p)t^2)(1-t^2)$.

The kernel functions corresponding to the values of $p = 1$, $p = 2$, $p = 3$, and $p = 4$ in Eq. (74) as given in Eqs.(75) through (78) are respectively the fourth-order hybrid Epanechnikov, biweight, triweight, and quadriweight kernels:

$$k^h_{[4,1]}(t) = \frac{3}{32}(27 - 70t^2 + 35t^4)k_{[2,0]}(t). \tag{75}$$

$$k^h_{[4,2]}(t) = \frac{5}{32}(13 - 42t^2 + 21t^4)k_{[2,1]}(t). \tag{76}$$

$$k^h_{[4,3]}(t) = \frac{7}{64}(17 - 66t^2 + 33t^4)k_{[2,2]}(t). \tag{77}$$

$$k^h_{[4,4]}(t) = \frac{9}{320}(63 - 286t^2 + 143t^4)k_{[2,3]}(t). \tag{78}$$

However, as $p \to \infty$, and by using Lemma 1.1, Eq. (74) tends to

$$k^h_{[4,\phi]}(t) = \begin{cases} \frac{1}{2}(6 - 7t^2 + t^4)k_{[2,\phi]}(t), & t \in (-\infty, \infty), \\ 0, & \text{elsewhere.} \end{cases} \tag{79}$$

Eq. (79) is thus the proposed fourth-order hybrid Gaussian kernel. Just as with the second-order hybrid kernels, the suggested fourth-order hybrid kernels, as presented in Eqs.(74) and (79), can be merged in the following manner:

$$k^h_{[4,p]}(t) = \begin{cases} \frac{\lambda}{8(p+1)}(\alpha(t,p) - \\ \qquad \beta(t,p))(1-t^2)^{p-1}, & p < \infty, \\ \frac{1}{2}(6 - t^2)(1-t^2)k_{[2,\phi]}(t), & p \to \infty, \end{cases} \tag{80}$$

where λ is as defined in Eq. (48). Eq.(80) can further be confirmed to adhere to the characteristics expected of fourth-order kernels, as outlined in Eq. (17). This is demonstrated by:

$$\int k(t)dt = \begin{cases} \frac{(2p+1)!!\sqrt{\pi}}{2^{p+1}\Gamma(p+\frac{3}{2})} = 1, & p < \infty, \\ 1, & p \to \infty. \end{cases} \tag{81}$$

$$\int tk(t)dt = \cdots = \int t^3k(t)dt = \begin{cases} 0, & p < \infty, \\ 0, & p \to \infty. \end{cases} \tag{82}$$

$$\int t^4k(t)dt = \begin{cases} \frac{3(2p+1)!!\sqrt{\pi}}{2^{p+2}(2p+7)\Gamma(p+\frac{5}{2})} \neq 0, & p < \infty, \\ 9 \neq 0, & p \to \infty. \end{cases} \tag{83}$$

Eqs.(81), (82), and (83), respectively, indicate that the proposed family of fourth-order hybrid polynomial kernels exhibits the following noteworthy properties: It functions as a PDF; all of its odd moments, including the second moment, are equal to zero; and it possesses a non-zero fourth moment. These characteristics align precisely with the stipulated criteria for higher-order kernels, as outlined in Eq. (17).

The graphs of the fourth-order classical beta polynomial kernels and the proposed fourth-order beta kernels for different values of p are presented in Figure 2(a) and (b). In the left panel (Figure 2(a)), the curves represent the shapes of fourth-order classical beta polynomial kernels, with each curve corresponding to a specific value of the parameter p. By analyzing these curves, we can observe the variations in the shape of the classical kernels as p changes. Like the seccond-order kernel, the curves exhibit a characteristic bell-shaped profile, which is a common feature of many kernel functions. With increasing values of p, the kernels become narrower and taller. This indicates that as p increases, the kernel concentrates its mass around the central point, making it more peaked. This visualization helps in understanding the impact of different p values on the kernel shape.

In the right panel (Figure 2(b)), the shapes of proposed fourth-order hybrid beta polynomial kernels are depicted. Similarly, each curve in this panel represents a particular value of p. By comparing the curves in Subfigure (b) to those in Subfigure (a), we can observe the differences in shape introduced by the hybrid approach. The distinct shapes exhibited by the hybrid kernels highlight the influence of the hybridization process on the behavior of the kernels.

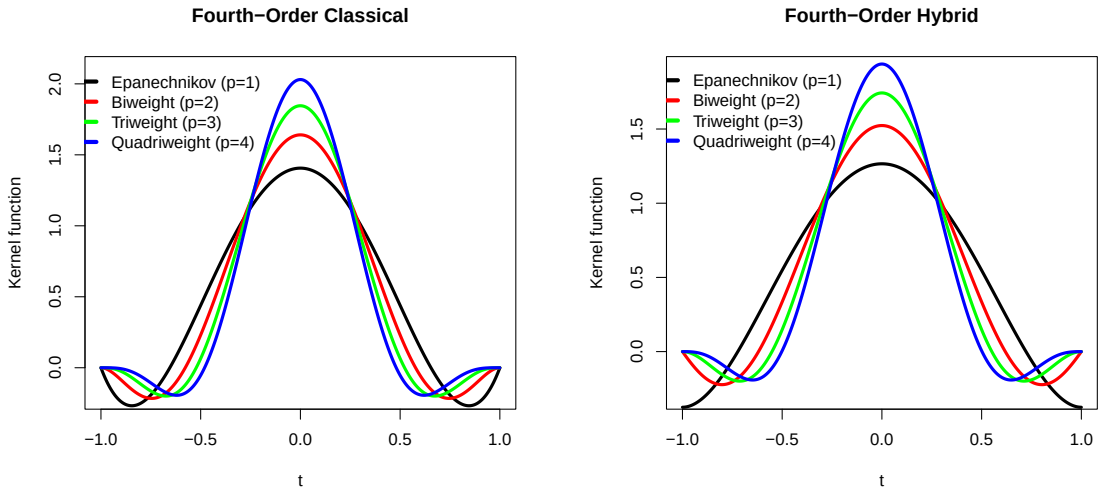
In a nutshell, Figure 2 provides a visual comparison between the shapes of fourth-order classical and hybrid beta polynomial kernels for different values of p. It allows for a clear comparison between the fourth-order classical and the proposed fourth-order hybrid kernels, showing potential differences in their behavior. This information can be valuable for selecting an appropriate kernel for a specific application or dataset. This comparison also aids in understanding how the hybridization process affects the kernel shape and is valuable for evaluating their suitability in various applications.

Table 1: The 4^{th} - Moments and Roughness of Classical and Hybrid Fourth-Order Beta Polynomial Kernels

Kernels	4^{th}-Moments		Roughness	
	Clas.	Hybr.	Clas.	Hybr.
Epanechnikov	$-\frac{1}{21}$	$-\frac{1}{15}$	$\frac{5}{4}$	$\frac{37}{32}$
Biweight	$-\frac{1}{33}$	$-\frac{3}{77}$	$\frac{805}{572}$	$\frac{3015}{2288}$
Triweight	$-\frac{3}{143}$	$-\frac{1}{39}$	$\frac{3780}{2431}$	$\frac{5215}{3536}$
Quadriweight	$-\frac{1}{65}$	$-\frac{1}{55}$	$\frac{28413}{16796}$	$\frac{74781}{46189}$

Table 1 presents the fourth moments and roughness values of classical and hybrid fourth-order beta polynomial kernels. The fourth moment provides a measure of the shape and symmetry of the kernel, while roughness represents the oscillatory behavior of the kernel. The values in the table reveal interesting insights about the characteristics of the kernels. The hybrid kernels consistently have more negative fourth moments compared to their classical counterparts. This indicates that the tails of the hybrid kernels spread out more than those of the classical kernels. However, the roughness values for the hybrid kernels are higher than those for the classical kernels across all kernel types. This implies that the hybrid kernels provide rougher density estimates, capturing more local variations in the data.

In a nutshell, the table shows that the proposed hybrid kernels tend to have less heavy tails and smoother shapes compared to their classical versions. This could have implications



(a) (b)

Figure 2: Plot of shapes of: (a) fourth-order classical beta polynomial kernels and (b) proposed fourth-order hybrid beta polynomial kernels for different values of p.

for applications where a smoother kernel with lighter tails is preferred. The specific numerical values provide a quantitative measure of these properties for each type of kernel.

4. Numerical Verification of Results

In this section, we will assess the performance of the proposed hybrid kernels in comparison to their classical kernel counterparts. To evaluate their effectiveness, we employ two numerical verification methods: Monte Carlo simulation experiments and real-life experiments using the AMISE measure in Theorem 2.2. Additionally, we compute the efficiencies of the kernels by using the generalized efficiency criterion in Eq. (33) to further analyze their performance. By employing these numerical verification methods, we aim to provide evidence and demonstrate the effectiveness of the proposed hybrid kernels in practical scenarios. The evaluation of their performance through Monte Carlo simulation experiments, real-life experiments, and efficiency computations will contribute to the validation of their superiority over the classical kernels.

4.1. Monte Carlo Experiment

To investigate the performance of the proposed fourth-order hybrid beta polynomial kernels in both small and large sample sizes, we conducted a series of simulation experiments. We employed a Monte Carlo approach with sample sizes of $n = 10$, $n = 25$, $n = 75$, and $n = 300$.

In these experiments, we considered a normal mixture density model defined as $X = \frac{1}{5}X_1 + \frac{1}{5}X_2 + \frac{3}{5}X_3$, where X_1, X_2, and X_3 follow normal distributions with means and variances $N(0, 1)$, $N\left(\frac{1}{4}, \frac{4}{9}\right)$, and $N\left(\frac{13}{12}, \frac{25}{81}\right)$ respectively.

To estimate the density, we generated a univariate random sample (X) and calculated its standard deviation. Then, we performed a repeated operation for $r = 1000$ iterations, using the AMISE in Theorem 2.2 to evaluate the performance. The average of the AMISE values (AMISE^*) is computed as:

$$\text{AMISE}^* = \frac{1}{r}\sum_{j=1}^{r}\text{AMISE}_j^{2m}. \tag{84}$$

where $m = 1, 2, \cdots$. To compute the average AMISE for the various classical and hybrid fourth-order kernels considered, we employed Eq. (84). By conducting these simulation experiments and calculating the AMISE, we can assess and compare the performance of the proposed hybrid beta polynomial kernels in different sample sizes. The results of these experiments are presented in Table 2, which includes the values for different simulated data sets.

Table 2: AMISE of Classical and Hybrid Fourth-Order Beta Polynomial Kernels

n	Kernels	Average AMISE	
		Classical	Hybrid
10	Epanechnikov	0.008915	0.008365
	Biweight	0.009874	0.009357
	Triweight	0.010754	0.010307
	Quadriweight	0.011550	0.011157
25	Epanechnikov	0.004438	0.004165
	Biweight	0.004915	0.004658
	Triweight	0.005352	0.005130
	Quadriweight	0.005747	0.005552
75	Epanechnikov	0.001515	0.001422
	Biweight	0.001677	0.001590
	Triweight	0.001826	0.001751
	Quadriweight	0.001961	0.001894
300	Epanechnikov	0.000387	0.000362
	Biweight	0.000427	0.000405
	Triweight	0.000465	0.000445
	Quadriweight	0.000499	0.000482

Table 2 compares the AMISE values between classical and hybrid fourth-order beta polynomial kernels for various sample sizes (n). Across all kernel types (Epanechnikov, biweight, triweight, and quadriweight) and sample sizes, hybrid kernels

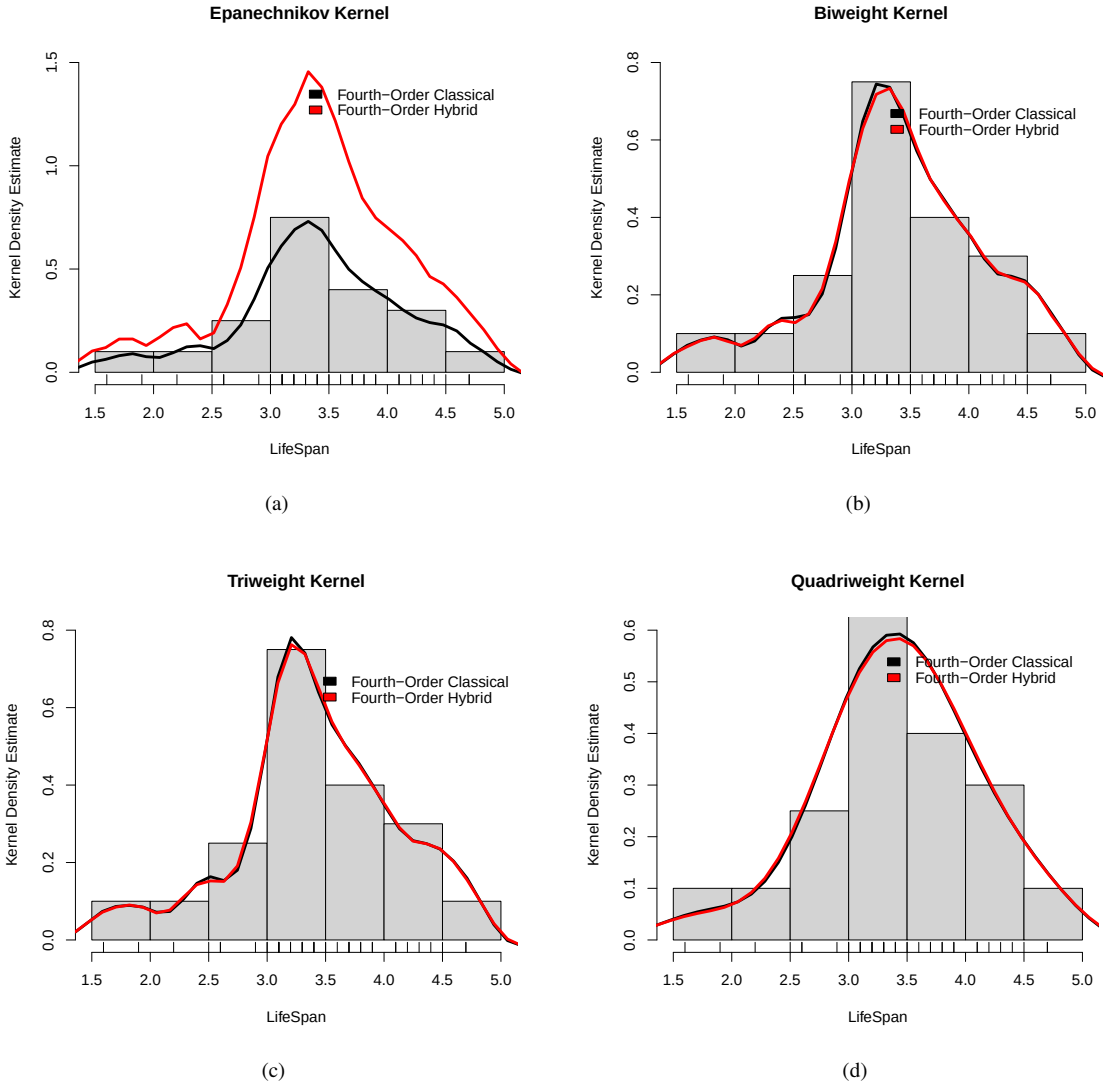


Figure 3: Histogram and kernel density estimates of: (a) fourth-order classical and hybrid Epanechnikov kernels; (b) fourth-order classical and hybrid Biweight kernels, (c) fourth-order classical and hybrid Triweight kernels and (d) fourth-order classical and hybrid quadriweight kernels using the lifespan of car battery data [24].

consistently exhibit lower AMISE values, indicating superior accuracy in density estimation.

Additionally, as sample size increases, AMISE values decrease, signifying improved accuracy with larger datasets. Overall, the table highlights the superior performance of hybrid kernels in achieving more accurate density estimates compared to classical kernels.

4.2. Real-Life Experiment

In this subsection, we assess the performance of the kernels embedded in the proposed fourth-order hybrid polynomial kernels' family using two real-life datasets. The first dataset represents the lifespan of car batteries in years, while the second dataset corresponds to the Old Faithful geyser eruptions. The first dataset is a unimodal asymmetrical dataset with a size of forty (40), obtained from [24], while the second dataset is a bimodal dataset with a size of two hundred and seventy-two (272), obtained from [33].

To evaluate the performance of the classical and hybrid fourth-order kernels on these real-life datasets, we constructed histograms for each dataset. Additionally, we overlaid the kernel density estimates using the four fourth-order kernels presented in this article. The resulting visualizations for the $n = 40$ and $n = 272$ datasets are presented in Figures 3 and 4, respectively. These data visualizations, analyses, and graphics were generated using Mathematica 11.3 and R Studio software.

Figure 3a shows the histogram and kernel density estimates for both fourth-order classical and hybrid Epanechnikov kernels. The histogram provides a visual representation of the distribution of lifespan data for car batteries, while the kernel density estimates provide smoothed probability density functions using either the classical or hybrid Epanechnikov kernels. These estimates serve as approximations of the under-

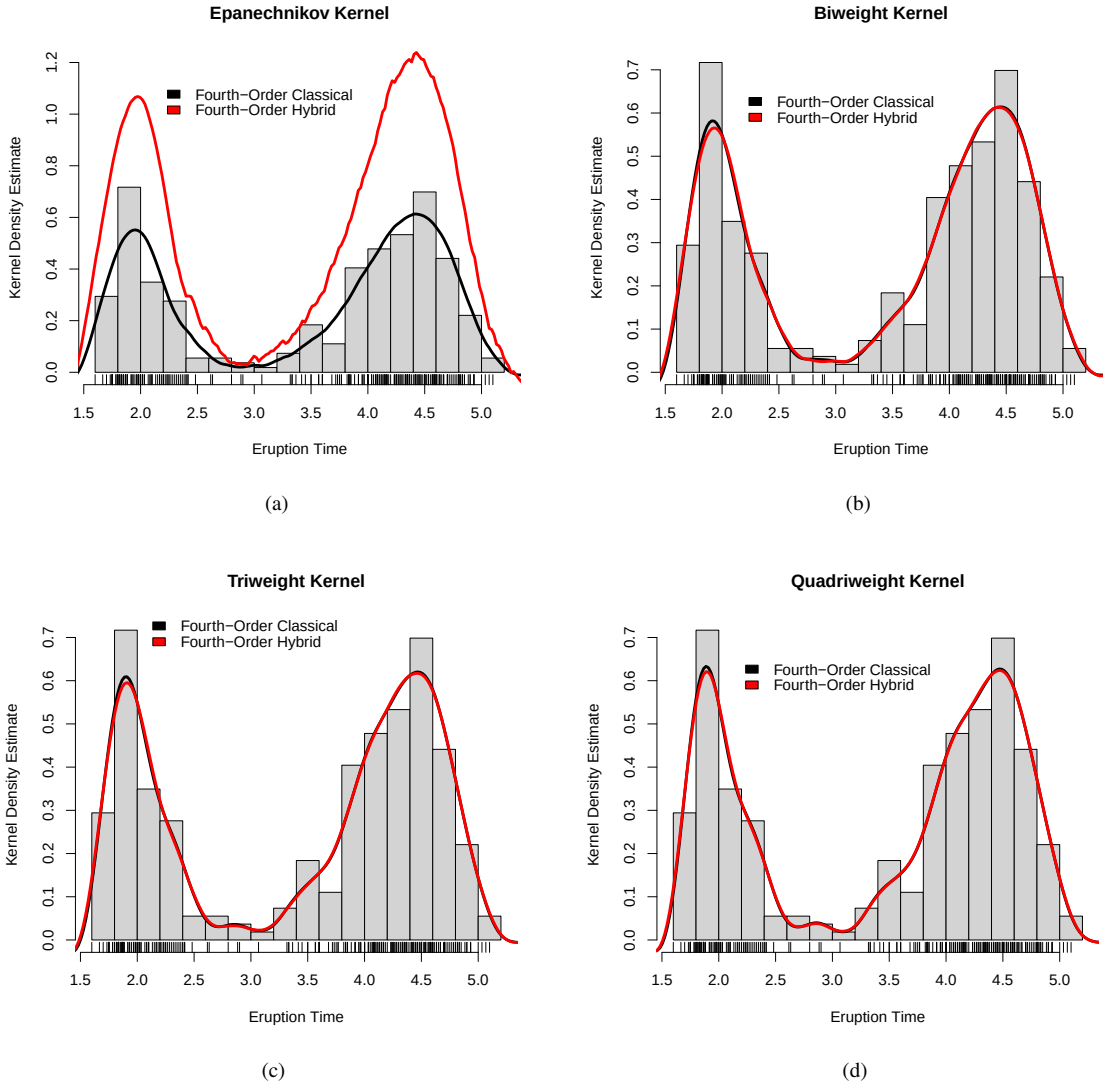


Figure 4: Histogram and kernel density estimates of: (a) fourth-order classical and hybrid Epanechnikov kernels; (b) fourth-order classical and hybrid Biweight kernels, (c) fourth-order classical and hybrid Triweight kernels and (d) fourth-order classical and hybrid Quadriweight kernels using Old Faithful eruption data [33].

lying probability density function of the data. Also, Figure 3b displays the histogram and kernel density estimates, but for fourth-order classical and hybrid biweight kernels. In addition, Figure 3c presents the histogram and kernel density estimates for fourth-order classical and hybrid triweight kernels, while Figure 3d showcases the histogram and kernel density estimates for fourth-order classical and hybrid quadriweight kernels.

In summary, Figure 3 assesses the suitability of the classical and hybrid kernels for modeling the car battery lifespan data. It can be seen that all the hybrid kernels fit the data better than their classical counterparts, most especially the Epanechnikov kernel. Thus, the quality of the fourth-order hybrid beta polynomial kernel density estimates is noteworthy, considering both their smoothness and accuracy. These kernel density estimates provide valuable insights into how well the kernels capture the distribution characteristics of the car battery lifespan data.

Like in Figure 3, we can equally assess the suitability of the classical and hybrid kernels for modeling the Old Faithful eruption data in Figure 4. This figure presents a detailed comparison of the histogram and kernel density estimates using different types of kernels on the Old Faithful eruption data. Figures 4a and 4b in the top panel, respectively, display the results for both fourth-order classical and hybrid Epanechnikov and biweight kernels. While in the bottom panel, Figures 4c and 4d, respectively, show the outcomes for fourth-order classical and hybrid triweight and quadriweight kernels.

It can be seen here that all the hybrid kernels fit the data better than their classical counterparts, most especially the Epanechnikov kernel. Hence, the quality of the fourth-order hybrid beta polynomial kernel density estimates, including their smoothness and accuracy, provides insights into how well the kernels capture the distribution patterns within the data.

Similar to the Monte Carlo study, we conducted a repeated operation for $r = 1000$ iterations using Eq. (84). This process allowed us to gather comprehensive results for the performance evaluation of the kernels. The obtained results, including per-

Table 3: AMISE of Classical and Hybrid Fourth-Order Beta Polynomial Kernels for real-life data.

n	Kernels	Average AMISE	
		Classical	Hybrid
40	Epanechnikov	0.003014	0.002829
	Biweight	0.003338	0.003164
	Triweight	0.003636	0.003485
	Quadriweight	0.003905	0.003772
272	Epanechnikov	0.000276	0.000259
	Biweight	0.000305	0.000289
	Triweight	0.000332	0.000318
	Quadriweight	0.000357	0.000344

Table 4: Efficiencies of Classical and Hybrid Fourth-Order Beta Polynomial Kernels.

Kernels	Efficiencies	
	Classical	Hybrid
Epanechnikov	1.0000 (100%)	1.0000 (100%)
Biweight	0.9969 (99.69%)	1.0020 (100.2%)
Triweight	0.9926 (99.26%)	0.9975 (99.75%)
Quadriweight	0.9889 (98.89%)	0.9935 (99.35%)

formance metrics and comparisons of the classical and hybrid fourth-order kernels, are presented in Table 3. This table serves as a summary of the analysis, presenting key findings from the evaluation process. Table 3 compares the AMISE for classical and hybrid fourth-order beta polynomial kernels applied to real-life data. For a sample size of 40, the hybrid kernels consistently outperform their classical counterparts, with improvements observed across all kernel types. The AMISE values for the hybrid kernels are lower compared to the classical kernels.

For a larger sample size of 272, similar trends are observed, with the hybrid kernels continuing to demonstrate superior performance in terms of lower AMISE values. In all, the hybrid kernels consistently provide more accurate density estimates compared to classical kernels, showcasing their effectiveness in real-life data scenarios.

4.3. Computation of Efficiencies

Table 4 presents the computation of the efficiencies of classical and hybrid fourth-order beta polynomial kernels. The efficiency values provide an indication of how well the different kernels perform in capturing the characteristics of the data. A higher efficiency value suggests a better fit to the underlying data distribution. Table 4 compares the efficiencies of classical and hybrid fourth-order beta polynomial kernels. Both types of Epanechnikov kernels achieve perfect efficiency (1.0000 or 100%), indicating optimal performance. The hybrid biweight and triweight kernels show slight improvements over their classical counterparts, with efficiencies of 1.002 (100.2%) and 0.9975 (99.75%), respectively. The hybrid quadriweight kernel also outperforms its classical counterpart with an efficiency of 0.9935 (99.35%).

This table clearly demonstrates that the hybrid approach enhances the performance of the kernels when compared to their classical counterparts. This improvement is evident as the efficiency values of the hybrid-biweight, hybrid-triweight, and hybrid-quadriweight kernels surpass those of their classical kernel counterparts.

5. Discussion of Results and Conclusion

In the field of statistics, a well-established principle dictates that as the error (AMISE) decreases, the estimator or kernel demonstrates higher accuracy. Consequently, a kernel function with a lower AMISE is deemed superior to one with a higher AMISE. This principle provides the foundation for our examination of Tables 2 and 3.

In addition to the aforementioned, the choice of kernel in relation to sample size is closely tied to the bias-variance trade-off. When dealing with a small sample size, employing a more complex kernel, such as a high-degree polynomial or a complex non-linear kernel, can lead to overfitting. This means that the model may become overly specialized for the training data, capturing noise and idiosyncrasies that fail to generalize well to unseen data. As the sample size increases (referred to as the medium sample size), the risk of overfitting diminishes. This makes higher-order kernels more valuable, as they can capture more intricate relationships in the data. In the case of large sample sizes, where there is an abundance of data to estimate complex relationships, higher-order kernels become increasingly advantageous. They enable the model to capture nuanced patterns that lower-order kernels might overlook [34, 35].

Table 2 furnishes average AMISE values for classical and hybrid fourth-order beta polynomial kernels in a Monte Carlo study. These values offer crucial insights into the accuracy of kernel density estimates, with lower AMISE values indicating superior fits to the underlying data distribution. Notably, the results consistently demonstrate the superior performance of the hybrid kernels over their classical counterparts in terms of AMISE values, regardless of sample size. This underscores the enhanced accuracy of the hybrid approach in estimating kernel densities. Additionally, as sample size increases, both classical and hybrid fourth-order kernels tend to achieve lower AMISE values, owing to the increased information available about the underlying distribution.

These findings are further corroborated by the analysis of real-life experiments presented in Table 3. In these experiments, both classical and hybrid kernels are evaluated across different sample sizes. The results consistently show that the hybrid kernels yield lower AMISE values compared to their classical counterparts, confirming the enhanced accuracy of the hybrid approach in estimating kernel densities for real-life datasets.

Turning to Table 4, the focus shifts to the computation of the efficiencies of classical and hybrid fourth-order beta polynomial kernels. Efficiency values offer insights into how well the kernels capture the characteristics of the data, with higher values indicating better fits to the underlying data distribution.

Table 4 provides a comparison of the efficiencies of classical and hybrid fourth-order beta polynomial kernels. Both classical and hybrid Epanechnikov kernels exhibit perfect efficiency, with a value of 1.0000 (100%). This indicates that both types of kernels perform optimally in capturing the characteristics of the data distribution.

The classical biweight kernel shows a high efficiency of 99.69%, indicating that it is very effective in capturing the characteristics of the data. On the other hand, the hybrid biweight

kernel demonstrates even higher efficiency, with a value of 1.0020 (100.2%). This suggests that the hybrid approach slightly outperforms the classical biweight kernel. Similar to the biweight kernels, the classical triweight kernel exhibits a high efficiency of 99.26%, indicating its effectiveness in capturing data characteristics. The hybrid triweight kernel also performs slightly better, with an efficiency of 99.75%. This again suggests a slight advantage for the hybrid approach.

The classical quadriweight kernel shows an efficiency of 98.89%, indicating it is effective but slightly less so compared to the other kernels. The hybrid quadriweight kernel demonstrates a higher efficiency of 99.35%, indicating that it is more effective in capturing the data characteristics compared to its classical counterpart.

This table highlights that, across various scenarios, the hybrid kernels generally exhibit a modest increase in efficiency compared to their classical counterparts. This observation underscores the potential advantages of employing the hybrid approach for density estimation, especially in situations where accurately capturing the characteristics of the data distribution is of paramount importance.

In conclusion, this article introduces a family of fourth-order hybrid beta polynomial kernels and evaluates their performance using average AMISE values and efficiency computations. The results from both Monte Carlo experiments and real-life experiments consistently demonstrate the superiority of the proposed hybrid kernels over their classical counterparts. The hybrid kernels exhibit improved accuracy and efficiency in estimating kernel densities, showcasing their effectiveness in capturing the characteristics of the data. These findings suggest that the hybrid approach holds promise for enhancing statistical analysis tasks compared to traditional kernels.

Acknowledgment

The author would like to express his sincere gratitude to the anonymous reviewers and the editorial board for their valuable feedback and constructive comments, which greatly contributed to the improvement of this article.

References

[1] W. S. Cleveland & R. McGill, "Graphical perception and graphical methods for analyzing scientific data", Science **229** (1985) 828. https://doi.org/10.1126/science.229.4716.828.

[2] E. R. Tufte, *The Visual display of quantitative information*, Graphics Press, Cheshire, 1983. https://doi.org/10.1002/bbpc.19850890519.

[3] E. Fix & J. L. Hodges, "Discriminatory analysis: nonparametric discrimination consistency properties ", (Report No. 4, Project No. **21.29.004**, USAF School of Aviation Medicine, Randolph Field, Texas, 1951 pp. 238. https://sci-hub.se/10.2307/1403797.

[4] H. Akaike, "An approximation to the density functions ", Annals of the Institute of Statistical Mathematics **6** (1954) 127. https://doi.org/10.1007/BF02900741.

[5] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function", Annals of Mathematical Statistics **27** (1956) 832. https://doi.org/10.1007/978-1-4419-8339-8-13.

[6] E. Parzen, "On the estimation of a probability density function and the mode", Annals of Mathematical Statistics **33** (1962) 1065. https://sci-hub.st/10.1214/aoms/1177704472.

[7] D. W. Scott & S. J. Sheather, "A reliable data-based bandwidth selection method for kernel density estimation", Journal of the Royal Statistical Society, Series B (Methodological) **53** (1991) 683. https://doi.org/10.1111/j.2517-6161.1991.tb01857.x.

[8] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, London, 1986. https://doi.org/10.1002/bimj.4710300745.

[9] A. W. Bowman & A. Azzalini, *Applied Smoothing Techniques for Data Analysis: The Kernels Approach with S-Plus Illustration*, Oxford University Press, UK (1997). https://doi.org/10.1093/oso/9780198523963.001.0001.

[10] D. W. Scott, *Multivariate Density Estimation: Theory, practice and visualization*, John Wiley & Sons Inc., New York, 1992. https://onlinelibrary.wiley.com/doi/book/10.1002/9780470316849.

[11] M. P. Wand & M. C. Jones, *Kernel Smoothing*, Chapman & Hall, London, 1995. https://doi.org/10.1201/b14876.

[12] P. Hall & J. S. Marron, "Lower bounds for bandwidth selection in density estimation", Probability Theory and Related Fields **90** (1991) 149. https://link.springer.com/article/10.1007/BF01192160.

[13] J. Simonoff, *Smoothing methods in statistics*, Springer, New York, 1996. https://link.springer.com/book/10.1007/978-1-4612-4026-6.

[14] A. Z. Zambom & R. Dias, "A review of kernel density estimation with applications to econometrics", arXiv:1212.2812 [stat.ME] **3** (2012) 1. https://doi.org/10.48550/arXiv.1212.2812.

[15] Y. Jeon & J. H. T. Kim, "A gamma kernel density estimation for insurance loss data", Insurance: Mathematics and Economics **53** (2013) 569. https://doi.org/10.1016/j.insmatheco.2013.08.009.

[16] X. Wang, M. Wahiduzzaman & A. Yeasmin, "A kernel density estimation approach and statistical generalized additive model of Western North Pacific Typhoon activities", Atmosphere **13** (2022) 1128. https://doi.org/10.3390/atmos13071128.

[17] J. C. C. Lopez & M. J. Sanz, "Improving kernels methods for density estimation in random differential equations problems", Mathematical and Computational Applications, **25** (2020) 33. https://doi.org/10.3390/mca25020033.

[18] J. E. Chacón, "Data-driven choice of the smoothing parametrization for kernel density estimators ", Canadian Journal of Statistics **37** (2009) 249. https://doi.org/10.1002/cjs.10016.

[19] V. A. Epanechnikov, "Nonparametric estimation of a multivariate probability density ", Theory of Probability and its Applications, **14** (1969) 153. https://doi.org/10.1137/1114019.

[20] B. A. Afere, *A family of multivariate higher-order hybrid polynomial kernels in kernel density estimation (Doctoral dissertation)*, University of Benin, Benin City, Nigeria, 2010.

[21] B. A. E. Afere, "On the convergence rate of d-dimensional fourth-order beta polynomial kernels", Recent Advances in Natural Sciences, **1** (2023) 29. https://doi.org/10.61298/rans.2023.1.2.29.

[22] F. Kimari, A. Adem & L. Kiti, "Efficiency of various bandwidth selection methods across different kernels ", IOSR Journal of Mathematics, **15** (2019) 55. https://doi.org/10.9790/5728-1503015562.

[23] M. G. Kendall & A. Stuart, *The Advanced Theory of Statistics*, Volume 2, Inference and Relationship, Griffin (1973). https://books.google.com.ng/books/about/The_Advanced_Theory_of_Statistics_Infere.html?id=ogTOAAAAMAAJ&redir_esc=y.

[24] C. C. Ishiekwene & B. A. E. Afere, "Higher-order window width selectors for emphirical data", Journal of Nigerian Statistical Association **14** (2001) 69.

[25] A. F. Ramsey, "An application of kernel density estimation via diffusion to group yield insurance", Paper prepared for presentation at the Agricultural & Applied Economics Association's (AAEA) Annual Meetings, Minneapolis, MN, https://ideas.repec.org/p/ags/aaea14/170173.html (2014).

[26] S. J. Shealter, "Density estimation", Statistical Science **19** (2004) 588. https://doi.org/10.1214/088342304000000297.

[27] B. A. Afere, "A new family of hybrid classical polynomial kernels in density estimation", International Journal of Research and Innovation in Applied Science (IJRIAS) **4** (2021) 145. https://www.rsisinternational.org/journals/ijrias/DigitalLibrary/Vol.6&Issue1/145-151.pdf.

[28] A. W. Bowman & P. J. Foster, "Adaptive smoothing and density-based tests of multivariate normality ", Journal of the American Statistical Association **88** (1993) 529. https://doi.org/10.2307/2290333.

[29] W. Hardle, M. Muller, S. Sperlich & A. Werwatz, *Nonparametric and Semiparametric models*, Springer-Verlag, Berlin, 2004. https://link.springer.com/book/10.1007/978-3-642-17146-8.

[30] M. C. Jones & P. J. Foster, "Generalized Jackknifing and higher-order kernels", Journal of Nonparametric Statistics **3** (1993) 81. https://doi.org/10.1080/10485259308832573.

[31] B. A. Afere & E. Alih, "On the reduction of global error of multivariate higher-order product polynomial kernels", Palestine Journal of Mathematics **8** (2019) 286. https://api.semanticscholar.org/CorpusID:209421718.

[32] B. E. Hansen, *Lecture note in nonparametric estimation [Lecture notes]*, University of Wisconsin, 2004. https://users.ssc.wisc.edu/~bhansen/718/NonParametrics1.pdf.

[33] GeyserTimes, *Eruption of old faithful geyser*, October 29 (2017) 13:55 [online database], Retrieved from https://geysertimes.org/data.php.

[34] J. Shawe-Taylor & N. Cristianini, *Kernel Methods for Pattern Analysis*, Cambridge University Press, UK (2004). https://vim.ustc.edu.cn/upload/article/files/d4/72/7d8483bd4f87-ae3f09be1e57664d/a5d0d3ff-de9b-4915-beda-1a2f44357ca2.pdf.

[35] T. Hastie, R. Tibshirani & J. Friedman, *The elements of statistical learning: data mining, inference, and prediction*, Springer, UK, 2009. https://link.springer.com/book/10.1007/978-0-387-84858-7.

Appendix: Symbols and acronyms

This appendix presents the symbols and acronyms used throughout this article.

- KDE = Kernel density estimation
- NKDE = Nonparametric kernel density estimation
- AMISE = Asymptotic mean integrated squared error
- MISE = Mean integrated squared error
- PDF = Probability density function
- $Eff(k_{sk})$ = Efficiency of symmetric kernel
- $C(k_{sk})$ = Constant associated with symmetric kernel
- $C(k_e)$ = Constant associated with Epanechnikov kernel
- $\int k^2(t)dt = L_2 -$ Norm
- $\int t^2 k(t)dt$ = Second moment or variance
- $k(t)$ = kernel function
- h = optimal bandwidth
- $k^h_{[2,p]}(t)$ = Second-order hybrid kernel
- $k^h_{[4,p]}(t)$ = Fourth-order hybrid kernel
- $k_{[2,p]}(t)$ = Second-order classical (traditional) kernel
- $k_{[4,p]}(t)$ = Fourth-order classical (traditional) kernel
- Clas = Classical kernel
- Hybr = hybrid kernel
- p = Power of the family of kernels