



On the cluster of the families of hybrid 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 study introduces a novel cluster of hybrid polynomial kernel families, designed to achieve significantly lower asymptotic mean integrated squared error compared to traditional kernels. These hybrid kernels are developed by heuristically combining classical polynomial kernels using probability axioms. An in-depth analysis of error propagation within these kernels is conducted, utilizing both simulation experiments and real-life datasets, including the Life Span of Batteries and COVID-19 datasets. The findings consistently demonstrate that the proposed hybrid kernels outperform their classical counterparts in various density estimation tasks across different distribution types and sample sizes. This research highlights the potential of hybrid polynomial kernels to enhance accuracy in density estimation, advocating for their adoption in statistical modelling and analysis.

DOI:10.46481/jnsps.2025.2044

Keywords: Kernel density estimation, Cluster of Hybrid kernels, Classical polynomial kernels, Global error, Monte Carlo simulation

Article History :

Received: 17 March 2024

Received in revised form: 12 August 2024

Accepted for publication: 09 September 2024

Available online: 22 October 2024

© 2025 The Author(s). Published by the [Nigerian Society of Physical Sciences](#) under the terms of the [Creative Commons Attribution 4.0 International license](#). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Communicated by: Joel Ndam

1. Introduction


Kernel density estimation (KDE) serves as a fundamental non-parametric technique for estimating the probability density function of a random variable based on a sample of data points. This subject area has been a cornerstone in mathematical statistics for the past seven decades, gaining prominence through the pioneering, albeit unpublished, works of Refs. [1, 2], as well as the published contributions of Fix & Hodges [3] and Akaike [4]. Since its inception, this field has seen remarkable progress. Notably, Silverman's monograph [5], as reported by Jiang &

Provost [6], has garnered references from over two thousand researchers, underscoring its pivotal role in advancing the domain.

Distinguished scholars such as those in Refs. [6–14], and [15] have made substantial and influential contributions, particularly in the field of bandwidth selection methods. Their collaborative endeavours have greatly enhanced the practical utility of KDE techniques.

A crucial aspect of KDE is the selection of the kernel function, which dictates the shape and smoothness of the resulting density estimate. The studies by Refs. [16, 17], as well as the references therein, have made noteworthy strides in the development of kernel functions. In these works, the developed kernel functions have demonstrated superior performance com-

*Corresponding author: Tel.: +234-913-539-8886.

Email address: baafere@gmail.com (Benson Ade Eniola Afere )

pared to the classical kernels found in the literature.

In KDE, classical kernel techniques such as Epanechnikov, biweight, triweight, and quadriweight are highly regarded due to their adaptability and efficiency in a range of applications [5]. In order to address dimensionality issues, recent advancements in KDE have brought in a hybrid nonparametric multivariate density estimator that combines kernel and exponential series approaches. Promising applications in financial risk management are demonstrated by this strategy, which improves computing feasibility and estimation accuracy [18]. Recent research [19, 20], which investigates hybrid polynomial kernels that combine the advantages of conventional (classical) kernel functions, reflects this growing interest. These hybrid methods provide an adaptable and flexible framework for better density estimation in a range of application scenarios.

This paper aims to expand upon the groundwork laid in Refs. [19, 20] by introducing a cluster of novel families of hybrid polynomial kernels within the realm of univariate kernel density estimation. The objective is to augment the toolkit available for practitioners in this field and offer a fresh perspective on the relationship between kernel functions and estimation accuracy in nonparametric density estimation. These families of hybrid kernels will provide practitioners with a versatile range of options for KDE. The selection of a specific family and its corresponding kernel function will have varying effects on smoothing and estimation accuracy, depending on the characteristics of the dataset under analysis.

The exploration of hybrid polynomial kernels is motivated by their potential to adapt to diverse data distributions and effectively capture intricate patterns. By conducting an in-depth examination of their mathematical formulations and theoretical properties, the study aims to provide insights into their applicability across different domains and highlight their comparative advantages over conventional kernel functions.

Empirical assessments are conducted to evaluate the performance of these hybrid polynomial kernels in both simulations and real-world scenarios. Comparative experiments and case studies demonstrate their effectiveness in handling various types of data, including unimodal and multimodal distributions, as well as skewed and heavy-tailed datasets. To further illustrate their relevance in nonparametric kernel density estimation (NKDE), these kernels are applied to real-world cases such as the lifespan of car batteries and the analysis of COVID-19 data. This contribution aims to enrich ongoing discussions and advancements in this fundamental statistical methodology, offering valuable insights and adding depth to the discourse.

The structure of the remaining article is organised as follows: Section 2 outlines the methodology. In Section 3, the proposed cluster of kernel families is introduced. Section 4 includes the simulation studies and real-life examples. Finally, Section 5 offers a discussion of the results and concludes with the findings.

2. Kernel density estimator

Given a random sample X_1, X_2, \dots, X_n of size n drawn independently and identically from a continuous distribution de-

scribed by the probability density function $f(x)$, the univariate KDE is formulated as follows:

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

In Equation (1), x signifies the range of observations, while $h > 0$ stands for the smoothing parameter, known by different names like bandwidth, bin width, or window width, as outlined by various authors, including [17]. Moreover, $K(\cdot)$ is denoted as the kernel function, which typically adheres to properties of symmetry and unimodality and abides by the following principles:

$$\left. \begin{array}{l} i. \quad \int K(t)dt = 1 \\ ii. \quad \int tK(t)dt = 0 \\ iii. \quad \int t^2K(t)dt = K_2 < \infty. \end{array} \right\} \quad (2)$$

Equation (2) reveals that every kernel function must satisfy the properties of a probability density function, implying that it must integrate to one (1). Additionally, its first moment about zero (mean) should integrate to zero (0), while its second moment about zero (variance) should integrate to a constant that is not equal to zero, [21].

As discussed earlier, Equation (1) was introduced to the mathematical statistics community through the works of Rosenblatt [3] and Parzen [4]. As a result, the estimator is often referred to as the Rosenblatt-Parzen estimator. Several error criteria are employed to evaluate the performance of KDE. These criteria can be categorised as pointwise error, global error, and asymptotic (approximate) error. Examples of pointwise error include mean squared error, root mean squared error, mean absolute error, etc. An example of a global error is the mean integrated squared error (MISE). The exact MISE is determined through convolution [22], while the asymptotic (approximate) MISE is derived using a Taylor series expansion, termed the asymptotic mean integrated squared error (AMISE).

Now, the expression for the mean integrated squared error of the estimator $\hat{f}_h(x)$, denoted as $(\text{MISE}_{\hat{f}_h(x)})$, is given by:

$$\text{MISE}_{\hat{f}_h(x)} = \text{E} \left(\int (\hat{f}_h(x) - f(x))^2 dx \right), \quad (3)$$

using necessary algebraic rules, Equation (3) becomes:

$$\text{MISE}_{\hat{f}_h(x)} = \int \text{Bias}^2 \hat{f}_h(x) dx + \int \text{Var} \hat{f}_h(x) dx, \quad (4)$$

where

$$\text{Bias} \hat{f}_h(x) = \text{E} \hat{f}_h(x) - f(x). \quad (5)$$

and

$$\text{Var} \hat{f}_h(x) = \text{E} \hat{f}_h^2(x) - \text{E}^2 \hat{f}_h(x). \quad (6)$$

Supposing the underlying density is suitably smooth and the second moment of the kernel is non-zero, the procedure involves substituting Equation (1) into Equations (5) and (6), followed by the utilisation of the univariate Taylor series expansion. This resulted in the transformation of Equations (5) and (6) into the following expressions, respectively:

$$\text{Bias} \hat{f}_h(x) = \frac{h^2}{2} K_2 f''(x) + o(h^2), \quad (7)$$

and

$$\text{Var} \hat{f}_h(x) = \frac{1}{nh} \|K\|_2^2 f(x) + o(nh)^{-1}. \quad (8)$$

where $\|K\|_2^2 = \int K^2(t)dt$ represents the L_2 -norm of the kernel function, which characterises the kernel function's degree of roughness. Additionally, $K_2 = \int t^2 K(t)dt$ in Equation (7) denotes the second moment of the kernel function, signifying the variance of the kernel function. For more details, refer to Ref. [5, 10]. By substituting Equations (7) and (8) into Equation (4) while neglecting higher-order terms, the expression for the AMISE can be derived as follows:

$$\text{AMISE} \hat{f}_h(x) = \frac{h^2}{2} K_2 \|f''(x)\|_2^2 + \frac{1}{nh} \|K\|_2^2 f(x), \quad (9)$$

where $\|f''(x)\|_2^2 = \int (f''(x))^2 dx$ represents the L_2 -norm of the unknown probability density function, it quantifies the degree of roughness in the function. This term essentially measures how rough the unknown probability density function is. The value of h in Equation (9) generally controls the balance between the two components of the AMISE. When the bias is large, the variance is reduced, and vice versa, illustrating the well-known "trade-off" between bias and variance. The optimal bandwidth is the smoothing parameter that minimises the AMISE. By differentiating Equation (9) with respect to h , the following result is obtained:

$$\frac{\delta \text{AMISE} \hat{f}_h(x)}{\delta h} = (K_2)^2 h^3 \|f''(x)\|_2^2 - \frac{\|K\|_2^2}{nh^2} = 0. \quad (10)$$

Solving the differential Equation (10) for h gives the optimal bandwidth as:

$$h_{\text{AMISE}} = \left(\frac{\|K\|_2^2}{(K_2)^2 \|f''(x)\|_2^2} \right)^{1/5} n^{-1/5}. \quad (11)$$

Substituting Equation (11) into Equation (9) gives:

$$\text{AMISE} \hat{f}_h(x) = 5 \cdot 4^{-4/5} \cdot (\|K\|_2^2)^{4/5} \times \left(\left(\frac{K_2}{2} \right)^2 \|f''(x)\| \right)^{1/5} \cdot n^{-4/5}. \quad (12)$$

Refer to Ref. [5, 10]. Utilising Equations (11) and (12), it can be deduced that the optimal bandwidth and the AMISE possess a convergence order of $o(n^{-1/5})$ and $o(n^{-4/5})$, respectively. However, in the context of the d -dimensional KDE, the convergence order of the optimal bandwidth and AMISE becomes $o(n^{-1/d+4})$ and $o(n^{-4/d+4})$, respectively, [20].

Furthermore, assuming the additional axioms as specified in Equation (2), if $\int t^{2m} K(t)dt = K_{2m} > 0$, Equations (11) and (12) respectively transform into:

$$h_{\text{AMISE}}^{2m} = \left(\frac{((2m)!)^2 \|K\|_2^2}{(4m) (K_{2m})^2 \|f^{(2m)}(x)\|_2^2} \right)^{\frac{1}{4m+1}} \times n^{-\frac{1}{4m+1}}, \quad m \in \mathbb{N}. \quad (13)$$

and

$$\text{AMISE}^{2m} \hat{f}_h(x) = (4m + 1) \times (4m)^{-\frac{4m}{4m+1}} \|K\|_2^{\frac{4m}{4m+1}} \times \left[\left(\frac{K_{2m}}{(2m)!} \right)^2 \|f^{(2m)}(x)\| \right]^{\frac{1}{4m+1}} \times n^{-\frac{4m}{4m+1}}, \quad m \in \mathbb{N}. \quad (14)$$

In Equations (13) and (14), n , K_{2m} , and $\|f^{(2m)}(x)\|$ represent the sample size, the $2m^{\text{th}}$ moment of any symmetric kernel function, and the L_{2m} -norm of the unknown probability density function, respectively. It is worth noting that the optimal bandwidth and the AMISE exhibit the order of convergence $o(n^{-\frac{1}{4m+1}})$ and $o(n^{-\frac{4m}{4m+1}})$, respectively.

3. Cluster of families of hybrid polynomial kernels

A cluster is characterised as a collection of similar entities arranged closely together. In Ref. [23], the concept of a cluster is further elucidated, describing it as "a group of items in which each item is 'close' to another item within the same group, while the members of distinct groups are 'distant' from each other." Therefore, this section elaborates on the establishment of the proposed cluster of kernel families. The foundational element employed here is the family of classical polynomial kernels, defined as follows:

$$K_c(t) = \{2^{2p+1} B(p+1, p+1)\}^{-1} (1-t^2)^p. \quad (15)$$

Here, p is referred to as the polynomial power, and $B(\cdot, \cdot)$ represents the beta density function. The kernel $K_c(t)$ has a support within the interval $[-1, 1]$. In Equation (15), distinct values of p result in different kernel functions. To illustrate, when $p = 0$, a uniform kernel is generated; when $p = 1$ and $p = 2$, a biweight kernel and a triweight kernel are produced, respectively. Remarkably, as the value of p approaches infinity ($p \rightarrow \infty$), $K_c(t)$ transforms into a Gaussian kernel.

Table 1. Classical polynomial kernels.

Value of p	Name of kernels	Function of kernels
$p = 0$	Uniform	$\frac{1}{2}$
$p = 1$	Epanechnikov	$\frac{3}{4}(1-t^2)$
$p = 2$	Biweight	$\frac{15}{16}(1-t^2)^2$
$p = 3$	Triweight	$\frac{35}{32}(1-t^2)^3$
$p = 4$	Quadriweight	$\frac{315}{256}(1-t^2)^4$

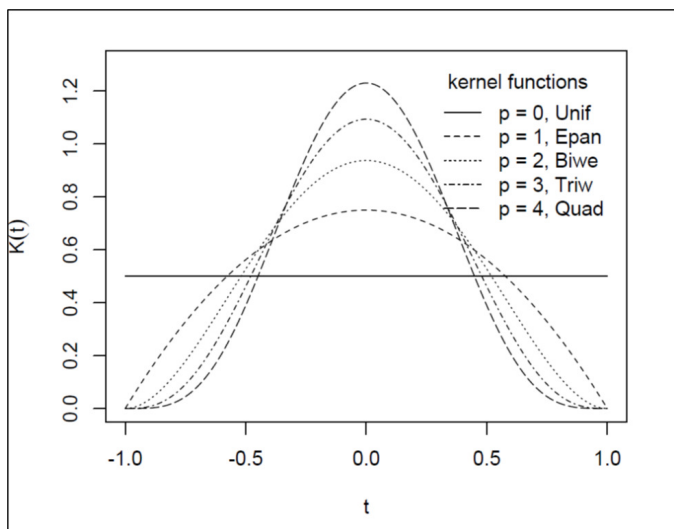


Figure 1. Plot of shapes of classical polynomial kernels for different values of p presented in Table 1.

Table 1 shows the kernel functions corresponding to various values of p . These kernels play a crucial role in smoothing data in kernel density estimation. The choice of kernel and its associated parameter (p) affect the smoothness and sensitivity of the estimated density curve. Hence, it's worth noting that as p increases, the kernel becomes more peaked at the centre and has heavier tails, which can lead to a smoother density estimate but may also result in over-smoothing if not chosen appropriately for the specific dataset.

Supplementing the information in Table 1, Figure 1 provides a clear visual representation of how the different values of p influence the shape and characteristics of the classical polynomial kernels. It visually demonstrates how the selection of p influences the weighting of data points in kernel density estimation, indicating that higher p values lead to more pronounced peaks at the centre and heavier tails. This visual representation is valuable for practitioners in selecting an appropriate kernel for their specific data and analysis needs. It allows them to make informed decisions about which kernel shape aligns best with the underlying distribution of their data.

As earlier said, this paper will extend the earlier papers [19] and [20]. Consider the simple formula:

$$K_H(t) = \rho_1 K_{[p-1]}(t) + \rho_2 K_{[p]}(t); \tag{16}$$

$$0 < \rho_1 < 1, \rho_1 + \rho_2 = 1.$$

where $K_{[p-1]}(t)$ and $K_{[p]}(t)$ represent the families of classical polynomial kernels of order $p-1$ and p , respectively, as defined in Equation (15). By substituting $\rho_1 = (i/10)$ into Equation (16) and replacing it with the corresponding values from Equation (15), along with the application of essential algebraic principles, the proposed cluster of hybrid polynomial kernel families

is established as follows:

$$K_{CL}(t) = \frac{2^{-p-2}}{5\Gamma(p+1)} [2ip(2p-1)!! \sum_{r=0}^{p-1} (-1)^r \binom{p-1}{r} + (10-i)(2p+1)!! \sum_{r=0}^p (-1)^r \binom{p}{r}] t^{2r}. \tag{17}$$

The kernel functions corresponding to individual families are displayed in Table 2, and it is evident that these kernel functions adhere to the properties of symmetric kernel functions as outlined in Equation (2).

Table 2 presents a cluster of hybrid classical polynomial kernels organised into distinct families indexed by i . These families encompass various kernels characterised by expressions rooted in the Epanechnikov, biweight, triweight, and quadriweight structures. Arranged in a matrix format, families are represented as rows and individual kernels as columns.

Crafted by blending traditional polynomial kernels with specific weights and coefficients, these hybrids exhibit a versatile nature, potentially designed for specific applications or tailored to capture diverse aspects of the underlying data distribution. Selection depends on the inherent characteristics of the data and the specific objectives of the analysis.

Hybrid classical polynomial kernels, featured in this table, showcase versatility by blending classical polynomial forms with additional parameters. This amalgamation allows them to effectively simulate intricate patterns in data and model complex data structures. The selection of a specific hybrid kernel is guided by nuanced analysis requirements, offering practitioners flexibility in addressing diverse dataset challenges. The table provides a comprehensive range of options for kernel density estimation, allowing researchers and practitioners to choose the most suitable hybrid kernel based on the characteristics of each family. In essence, it serves as a valuable reference for statisticians and data analysts seeking adaptable tools for their analytical needs.

Analytical validation will demonstrate that the kernels included in Equation (17), individually described in Table 2, indeed adhere to the properties outlined in Equation (2). Focussing initially on the first family (F1), the outcomes for each kernel in the cluster of kernel families are presented, as illustrated in Table 3.

For Epanchnikov kernel:

(i)

$$\int K(t) dt = \int_{-1}^1 \frac{1}{10} (29 - 27t^2) dt$$

$$= \frac{1}{40} \left(29t - \frac{27t^3}{3} \right) \Big|_{-1}^1$$

$$= \frac{1}{40} (40) = 1.$$

(ii)

$$\int tK(t) dt = \int_{-1}^1 \frac{1}{10} t(29 - 27t^2) dt$$

Table 2. A cluster of families of hybrid classical polynomial kernels

Family (<i>i</i>)	Kernels			
	Epanechnikov	Biweight	Triweight	Quadriweight
F1	$\frac{1}{40}(29 - 27t^2)$	$\frac{3}{160}(49 - 94t^2 + 45t^4)$	$\frac{3}{64}(1 - t^2)^2(23 - 21t^2)$	$\frac{7}{512}(1 - t^2)^3(89 - 81t^2)$
F2	$\frac{1}{10}(7 - 6t^2)$	$\frac{3}{20}(6 - 11t^2 + 5t^4)$	$\frac{1}{16}(1 - t^2)^2(17 - 14t^2)$	$\frac{7}{64}(1 - t^2)^3(11 - 9t^2)$
F3	$\frac{3}{40}(9 - 7t^2)$	$\frac{3}{160}(47 - 82t^2 + 35t^4)$	$\frac{1}{64}(1 - t^2)^2(67 - 49t^2)$	$\frac{21}{512}(1 - t^2)^3(29 - 21t^2)$
F4	$\frac{1}{20}(13 - 9t^2)$	$\frac{3}{80}(23 - 38t^2 + 15t^4)$	$\frac{3}{32}(1 - t^2)^2(11 - 7t^2)$	$\frac{7}{256}(1 - t^2)^3(43 - 27t^2)$
F5	$\frac{1}{8}(5 - 3t^2)$	$\frac{3}{32}(9 - 14t^2 + 5t^4)$	$\frac{3}{64}(1 - t^2)^2(13 - 7t^2)$	$\frac{35}{512}(1 - t^2)^3(17 - 9t^2)$
F6	$\frac{1}{40}(23 - 9t^2)$	$\frac{3}{160}(43 - 58t^2 + 15t^4)$	$\frac{21}{64}(1 - t^2)^2(3 - t^2)$	$\frac{7}{512}(1 - t^2)^3(83 - 27t^2)$
F7	$\frac{1}{40}(23 - 9t^2)$	$\frac{3}{40}(11 - 16t^2 + 5t^4)$	$\frac{1}{16}(1 - t^2)^2(16 - 7t^2)$	$\frac{21}{128}(1 - t^2)^3(7 - 3t^2)$
F8	$\frac{1}{20}(11 - 3t^2)$	$\frac{3}{80}(21 - 26t^2 + 5t^4)$	$\frac{1}{32}(1 - t^2)^2(31 - 7t^2)$	$\frac{7}{256}(1 - t^2)^3(41 - 9t^2)$
F9	$\frac{3}{40}(7 - t^2)$	$\frac{3}{160}(41 - 46t^2 + 5t^4)$	$\frac{1}{64}(1 - t^2)^2(61 - 7t^2)$	$\frac{63}{512}(1 - t^2)^3(9 - t^2)$

$$= \frac{1}{40} \left(\frac{29t^2}{2} - \frac{27t^4}{4} \right) \Big|_{-1}^1$$

$$= \frac{1}{40}(0) = 0.$$

(iii)

$$\int t^2 K(t) dt = \int_{-1}^1 \frac{1}{10} t^2 (29 - 27t^2) dt$$

$$= \frac{1}{40} \left(\frac{29t^3}{3} - \frac{27t^5}{5} \right) \Big|_{-1}^1$$

$$= \frac{16}{75} \neq 0.$$

Hence, F1-Biweight is a symmetric kernel.

For triweight kernel

(i)

$$\int K(t) dt = \int_{-1}^1 \frac{3}{64} (1 - t^2)^2 (23 - 21t^2) dt$$

$$= \frac{3}{64} \left(23t - \frac{67t^3}{3} + \frac{65t^5}{5} - \frac{21t^7}{7} \right) \Big|_{-1}^1$$

$$= 1.$$

(ii)

Hence, F1-Epanechnikov is a symmetric kernel.

For biweight kernel

(i)

$$\int K(t) dt = \int_{-1}^1 \frac{3}{160} (1 - t^2)(49 - 45t^2) dt$$

$$= \frac{3}{160} \left(49t - \frac{94t^3}{3} + \frac{45t^5}{5} \right) \Big|_{-1}^1$$

$$= \frac{3}{160} \left(\frac{160}{3} \right) = 1.$$

(ii)

$$\int tK(t) dt = \int_{-1}^1 \frac{3}{160} t(1 - t^2)(49 - 45t^2) dt$$

$$= \frac{3}{160} \left(\frac{49t^2}{2} - \frac{94t^4}{4} + \frac{45t^6}{6} \right) \Big|_{-1}^1$$

$$= \frac{3}{160}(0) = 0.$$

(iii)

$$\int t^2 K(t) dt = \int_{-1}^1 \frac{3}{160} t^2 (1 - t^2)(49 - 45t^2) dt$$

$$= \frac{3}{160} \left(\frac{49t^3}{3} - \frac{94t^5}{5} + \frac{45t^7}{7} \right) \Big|_{-1}^1$$

$$= \frac{53}{280} \neq 0.$$

$$\int tK(t) dt = \int_{-1}^1 \frac{3}{64} t(1 - t^2)^2 (23 - 21t^2) dt$$

$$= \frac{3}{64} \left(\frac{23t^2}{2} - \frac{67t^4}{4} + \frac{65t^6}{6} - \frac{21t^8}{8} \right) \Big|_{-1}^1$$

$$= 0.$$

(iii)

$$\int t^2 K(t) dt = \int_{-1}^1 \frac{3}{64} t^2 (1 - t^2)^2 (23 - 21t^2) dt$$

$$= \frac{3}{64} \left(\frac{23t^3}{3} - \frac{67t^5}{5} + \frac{65t^7}{7} - \frac{21t^9}{9} \right) \Big|_{-1}^1$$

$$= \frac{4}{35} \neq 0.$$

Hence, F1-Triweight is a symmetric kernel.

For quadriweight kernel

(i)

$$\int K(t) dt = \int_{-1}^1 \frac{7}{512} (1 - t^2)^3 (89 - 81t^2) dt$$

$$= \frac{7}{512} \times \frac{512}{7}$$

$$= 1.$$

(ii)

$$\begin{aligned} \int tK(t) dt &= \int_{-1}^1 \frac{7}{512} t(1-t^2)^3(89-81t^2) dt \\ &= \frac{7}{512} (0) \\ &= 0. \end{aligned}$$

(iii)

$$\begin{aligned} \int t^2K(t) dt &= \int_{-1}^1 \frac{7}{512} t^2(1-t^2)^3(89-81t^2) dt \\ &= \frac{7}{512} \times \frac{23552}{10395} \\ &= \frac{46}{495}. \end{aligned}$$

Therefore, F1-Quadriweight is categorised as a symmetric kernel, affirming that the kernels within the F1 family adhere to the properties outlined in Equation (2).

Table 3 provides a summary of key properties for a range of families of hybrid classical polynomial kernels. Each row corresponds to a specific family labelled from F1 to F9, and each column pertains to a particular property or moment of the respective kernel. The first property holds true as the kernel function integrates to unity, signifying consistent total probability coverage across all families. Furthermore, the first moment, yielding a value of zero, implies the distribution is centred around $t = 0$. Lastly, the second moment (variance), an indicator of skewness, exhibits finite values that differ among families, showcasing variations in kernel asymmetry. This fulfils the previously mentioned properties outlined in Equation (2), and thus, consequently, the kernels presented in Table 2 can be classified as symmetric kernels.

4. Data visualisations and numerical experiments

In this section, the discussion will be anchored in the visualisation of data and the examination of numerical results, with the aim of assessing the performance of the proposed kernel functions. This evaluation will be conducted using a combination of simulated datasets through a Monte Carlo experiment and real-life datasets.

4.1. Monte carlo experiments

Monte Carlo experiments were conducted to investigate four different mixture densities, each representing distinct distributional characteristics. The sample sizes used in the experiments were $n = 10$, $n = 25$, $n = 300$, and $n = 1500$. The mixture densities studied are as follows:

1. Unimodal Distribution (X_1):

$$X_1 = \frac{3}{5}X_{11} + \frac{2}{5}X_{12}.$$

where X_{11} follows a normal distribution with mean 0 and variance 1 ($X_{11} \sim N(0, 1)$), and X_{12} follows a normal distribution with mean 2 and variance 1 ($X_{12} \sim N(2, 1)$).

2. Multimodal Distribution (X_2):

$$X_2 = \frac{1}{3}X_{21} + \frac{1}{3}X_{22} + \frac{1}{3}X_{23}.$$

where X_{21} follows a normal distribution with mean -2 and variance 1 ($X_{21} \sim N(-2, 1)$), X_{22} follows a normal distribution with mean 0 and variance 1 ($X_{22} \sim N(0, 1)$), and X_{23} follows a normal distribution with mean 2 and variance 1 ($X_{23} \sim N(2, 1)$).

3. Skewed Distribution (X_3):

$$X_3 = \frac{3}{10}X_{31} + \frac{7}{10}X_{32}.$$

where X_{31} follows a normal distribution with mean -1 and variance 1 ($X_{31} \sim N(-1, 1)$), and X_{32} follows a normal distribution with mean 2 and variance 1 ($X_{32} \sim N(2, 1)$).

4. Heavy-Tailed Distribution (X_4):

$$X_4 = \frac{4}{5}X_{41} + \frac{1}{5}X_{42}.$$

where X_{41} follows a normal distribution with mean 0 and variance 1 ($X_{41} \sim N(0, 1)$), and X_{42} follows a Cauchy distribution with location parameter 0 and scale parameter 1 ($X_{42} \sim \text{Cauchy}(0, 1)$).

The examination of various distributions in Table 4 reveals unique traits that are essential to comprehending their behaviours. High variance and excessive kurtosis characterise a heavy-tailed distribution, suggesting the existence of notable outliers and prominent tails. The Cauchy component has a significant impact on this distribution, which adds to its wide dispersion and high skewness. These characteristics point to a distribution with significant tail effects and wide standard deviations.

The multimodal distribution, on the other hand, shows nearly normal kurtosis values and a low variance. Its skewness is relatively negligible, indicating a multi-peak, symmetrical distribution. This is consistent with the hypothesis that a distribution with many unique modes would result from mixing multiple normal distributions with various means.

In contrast, the skewed distribution has a closely aligned mean and median, suggesting a slight skewness. This distribution exhibits a somewhat asymmetric shape with a modest variance and kurtosis close to normal values. The moderate variance points to a distribution that is less extreme than the heavy-tailed one, while the balanced mean and median show that the skewness is not too great.

The kurtosis and skewness values of the unimodal distribution are modest and resemble those of a normal distribution. An example of a typical unimodal form is seen in the moderate variance and almost similar mean and median. The properties of this distribution confirm that it is single-peaked, as it closely resembles a normal distribution.

Within each of the analysed density models, univariate random variables (X) are generated, and subsequently, the standard deviation parameters are estimated based on these generated values. Following this step, the AMISE expression denoted by Equation (14) is employed. A comprehensive simulation

Table 3. Highlight of the properties of cluster of families of hybrid classical polynomial kernels in Equation (17).

Family (i)	Kernels											
	Epanechnikov			Biweight			Triweight			Quadriweight		
	$\int K(t)dt$	$\int tK(t)dt$	$\int t^2K(t)dt$	$\int K(t)dt$	$\int tK(t)dt$	$\int t^2K(t)dt$	$\int K(t)dt$	$\int tK(t)dt$	$\int t^2K(t)dt$	$\int K(t)dt$	$\int tK(t)dt$	$\int t^2K(t)dt$
F1	1	0	$\frac{16}{75}$	1	0	$\frac{26}{175}$	1	0	$\frac{4}{35}$	1	0	$\frac{46}{495}$
F2	1	0	$\frac{17}{75}$	1	0	$\frac{27}{175}$	1	0	$\frac{37}{315}$	1	0	$\frac{47}{495}$
F3	1	0	$\frac{6}{25}$	1	0	$\frac{4}{25}$	1	0	$\frac{38}{315}$	1	0	$\frac{16}{495}$
F4	1	0	$\frac{19}{75}$	1	0	$\frac{29}{175}$	1	0	$\frac{13}{315}$	1	0	$\frac{165}{49}$
F5	1	0	$\frac{4}{15}$	1	0	$\frac{6}{35}$	1	0	$\frac{105}{63}$	1	0	$\frac{495}{99}$
F6	1	0	$\frac{7}{25}$	1	0	$\frac{31}{175}$	1	0	$\frac{41}{315}$	1	0	$\frac{17}{165}$
F7	1	0	$\frac{22}{75}$	1	0	$\frac{32}{175}$	1	0	$\frac{2}{15}$	1	0	$\frac{32}{495}$
F8	1	0	$\frac{23}{75}$	1	0	$\frac{33}{175}$	1	0	$\frac{43}{315}$	1	0	$\frac{53}{495}$
F9	1	0	$\frac{8}{25}$	1	0	$\frac{34}{175}$	1	0	$\frac{44}{315}$	1	0	$\frac{6}{55}$

Table 4. Exploratory data analysis

Distribution	Mean	Median	Variance	Skewness	Kurtosis
Heavy-Tailed	0.209	0.0168	120.0	35.3	1768.0
Multimodal	0.000920	0.00358	0.333	-0.0323	2.95
Skewed	1.10	1.10	0.575	-0.0368	3.10
Unimodal	0.795	0.792	0.522	0.0169	2.93

is conducted encompassing $r = 1000$ runs, ultimately yielding the average AMISE denoted as $AMISE^*$.

$$AMISE^* = \frac{1}{r} \sum_{j=1}^r AMISE_j^{2m}, m \in \mathbb{N}. \tag{18}$$

Equation (18) was computed for hybrid Epanechnikov, biweight, triweight, and quadriweight kernels for all the univariate densities as presented in Tables 5, 6, 7, and 8 for four sample sizes of unimodal, multimodal, skewed, and heavy-tailed distributions, respectively.

Table 5 provides a comparison between the average global errors of the hybrid kernel family and the classical kernel family for four different sample sizes of a unimodal distribution. A lower average global error is regularly observed for hybrid kernels than for classical kernels with a sample size of $n = 10$. Hybrid kernels continue to have a decreased error rate, and this trend continues through the sample size of $n = 25$. Both classical and hybrid kernels attain very low average global errors when sample sizes approach $n = 300$ and $n = 1500$. However, hybrid kernels persistently outperform classical ones, confirming their efficacy in lowering error across a range of sample sizes.

The average global errors for a multimodal distribution for four sample sizes comparing hybrid and classical kernels are shown in Table 6. Hybrid kernels regularly show lower average global errors than classical kernels, with an advantage that becomes more substantial at smaller sample numbers, as seen in Table 5. Overall mistakes reduce as sample size grows, but hybrid kernels still show a small accuracy advantage, with the F9 kernel consistently exhibiting the highest performance.

The average global errors of the classical and hybrid kernels with a skewed distribution are compared for four sample sizes

(10, 25, 300, and 1500) in Table 7. Classical kernels at $n = 10$ vary from 0.00277444 to 0.00398071, but hybrid kernels with F9 at 0.00237631 have smaller errors. Errors reduce for all kernels as the sample size grows. F9 achieves the lowest error at 0.00001433, while classical kernels range from 0.00001663 to 0.00002364 by $n = 1500$. With hybrid kernels continuously outperforming classical ones, F9 in particular, which exhibits the highest performance across all sample sizes, the error reduction becomes more noticeable in larger sample sizes.

For a heavy-tailed distribution, Table 8 presents a comparison between classical and hybrid kernels for four sample sizes (10, 25, 300, and 1500). As sample sizes increase, errors reduce, and hybrid kernels routinely outperform classical ones. Hybrid F9 scores 0.00296317 at $n = 10$, while classical errors range from 0.00345962 to 0.00496379. Classical errors decrease to 0.00000265 – 0.00000377 by $n = 1500$; F9 achieves the lowest error reduction at 0.00000229, indicating the greater error reduction of hybrid kernels as sample sizes rise.

4.2. Real-life experiments

In this section, we delve into evaluating the performance of kernels integrated within our proposed cluster of kernel families, employing two real-life datasets. The first dataset comprises unimodal data, consisting of forty (40) observations detailing the lifespan of car batteries in years [24]. On the other hand, the second dataset contains bimodal data, consisting of a sample size of two hundred and fourteen (214) generated from the density derived from the COVID-19 data, as extracted from Osatohanmwun *et al* [25]. The kernel density estimations illustrating the classical kernels of these two datasets, as well as the superimposed kernel density estimates for all the families for each kernel function, are visually presented in Figures 2 (a, b,

Table 5. Comparison of average global error of kernels in the cluster of families of hybrid kernels with the kernels in the family of classical kernels for four sample sizes of unimodal distribution.

		AMISE*			
<i>n</i>	Family <i>i</i>)	Epanechnikov	Biweight	Triweight	Quadriweight
10	Classical	0.00215145	0.00250927	0.00281777	0.00308686
	F1	0.00210128	0.00247561	0.00279221	0.00306615
	F2	0.00204767	0.00243530	0.00276010	0.00303921
	F3	0.00199888	0.00239494	0.00272712	0.00301114
	F4	0.00195606	0.00235550	0.00269418	0.00298277
	F5	0.00191970	0.00231737	0.00266163	0.00295445
	F6	0.00189004	0.00228076	0.00262965	0.00292635
	F7	0.00186726	0.00224579	0.00259836	0.00289857
	F8	0.00185146	0.00221255	0.00256784	0.00287118
	F9	0.00184272	0.00218109	0.00253812	0.00284424
25	Classical	0.00138092	0.00160964	0.00180682	0.00197877
	F1	0.00134887	0.00158815	0.00179050	0.00196555
	F2	0.00131462	0.00156240	0.00176999	0.00194834
	F3	0.00128343	0.00153660	0.00174892	0.00193041
	F4	0.00125607	0.00151139	0.00172787	0.00191229
	F5	0.00123284	0.00148702	0.00170706	0.00189419
	F6	0.00121390	0.00146361	0.00168662	0.00187623
	F7	0.00119937	0.00144124	0.00166661	0.00185847
	F8	0.00118931	0.00141999	0.00164710	0.00184097
	F9	0.00118379	0.00139987	0.00162810	0.00182374
300	Classical	0.00010932	0.00012718	0.00014257	0.00015598
	F1	0.00010683	0.00012551	0.00014130	0.00015496
	F2	0.00010416	0.00012350	0.00013970	0.00015361
	F3	0.00010172	0.00012149	0.00013806	0.00015222
	F4	0.00009959	0.00011952	0.00013642	0.00015080
	F5	0.00009778	0.00011762	0.00013479	0.00014939
	F6	0.00009631	0.00011579	0.00013320	0.00014800
	F7	0.00009518	0.00011404	0.00013163	0.00014660
	F8	0.00009441	0.00011238	0.00013011	0.00014524
	F9	0.00009399	0.00011080	0.00012862	0.00014389
1500	Classical	0.00002053	0.00002384	0.00002670	0.00002918
	F1	0.00002007	0.00002353	0.00002646	0.00002899
	F2	0.00001957	0.00002316	0.00002617	0.00002875
	F3	0.00001912	0.00002279	0.00002586	0.00002849
	F4	0.00001872	0.00002242	0.00002556	0.00002822
	F5	0.00001839	0.00002207	0.00002526	0.00002796
	F6	0.00001812	0.00002173	0.00002496	0.00002770
	F7	0.00001791	0.00002140	0.00002467	0.00002745
	F8	0.00001777	0.00002110	0.00002439	0.00002719
	F9	0.00001769	0.00002080	0.00002411	0.00002694

c, d) and 3 (a, b, c, d), respectively, for the lifespan data and COVID-19 data. It is noteworthy that all aspects of data visualisation, analysis, and graphical representation were conducted using Mathematica 11.3 and R Studio software. The outcomes of these analyses are meticulously compiled within Tables 5 through 10.

The figure (Figure 2) comprises four subplots, each depicting the density estimation performance of different types of kernels across various families, including classical kernels, using the lifespan of car batteries in years as the data source. In

the top-left panel, the plot showcases the classical Epanechnikov kernel, while subplots F1 to F9 display density estimations for various families within the Epanechnikov hybrid kernel. Typically, the Epanechnikov kernel generates density estimates characterised by a triangular shape with a flattened peak, facilitated by a high bandwidth value of 163.85, resulting in smooth density curves.

Moving to the top-right panel, the plots depict the hybrid biweight kernel for all the families within the cluster of kernel families, alongside the classical biweight kernel. Renowned

Table 6. Comparison of average global error of kernels in the cluster of families of hybrid kernels with the kernels in the family of classical kernels for four samples sizes of multimodal distribution.

		AMISE*			
<i>n</i>	Family	Epanechnikov	Biweight	Triweight	Quadriweight
10	Classical	0.00171855	0.00200437	0.00225080	0.00246575
	F1	0.00167848	0.00197748	0.00223038	0.00244920
	F2	0.00163565	0.00194528	0.00220473	0.00242768
	F3	0.00159668	0.00191304	0.00217839	0.00240525
	F4	0.00156247	0.00188154	0.00215208	0.00238260
	F5	0.00153343	0.00185108	0.00212607	0.00235998
	F6	0.00150974	0.00182184	0.00210053	0.00233753
	F7	0.00149154	0.00179391	0.00207554	0.00231534
	F8	0.00147892	0.00176735	0.00205115	0.00229346
	F9	0.00147194	0.00174222	0.00202742	0.00227194
25	Classical	0.00076253	0.00088883	0.00099771	0.00109266
	F1	0.00074484	0.00087696	0.00098870	0.00108536
	F2	0.00072592	0.00086274	0.00097737	0.00107586
	F3	0.00070870	0.00084850	0.00096574	0.00106596
	F4	0.00069359	0.00083458	0.00095411	0.00105595
	F5	0.00068076	0.00082112	0.00094262	0.00104596
	F6	0.00067031	0.00080819	0.00093134	0.00103604
	F7	0.00066228	0.00079584	0.00092029	0.00102623
	F8	0.00065673	0.00078410	0.00090951	0.00101657
	F9	0.00065368	0.00077300	0.00089902	0.00100705
300	Classical	0.00006721	0.00007819	0.00008765	0.00009590
	F1	0.00006568	0.00007716	0.00008687	0.00009527
	F2	0.00006404	0.00007593	0.00008589	0.00009444
	F3	0.00006254	0.00007469	0.00008488	0.00009358
	F4	0.00006123	0.00007348	0.00008387	0.00009271
	F5	0.00006011	0.00007231	0.00008287	0.00009185
	F6	0.00005921	0.00007119	0.00008189	0.00009098
	F7	0.00005852	0.00007011	0.00008093	0.00009013
	F8	0.00005804	0.00006909	0.00007999	0.00008929
	F9	0.00005779	0.00006812	0.00007908	0.00008847
1500	Classical	0.00001499	0.00001741	0.00001949	0.00002131
	F1	0.00001465	0.00001718	0.00001932	0.00002117
	F2	0.00001429	0.00001691	0.00001910	0.00002099
	F3	0.00001396	0.00001664	0.00001888	0.00002080
	F4	0.00001367	0.00001637	0.00001866	0.00002061
	F5	0.00001343	0.00001611	0.00001844	0.00002042
	F6	0.00001323	0.00001586	0.00001822	0.00002023
	F7	0.00001308	0.00001563	0.00001801	0.00002004
	F8	0.00001297	0.00001540	0.00001780	0.00001985
	F9	0.00001292	0.00001519	0.00001760	0.00001967

for its robustness to outliers, the biweight kernel yields density estimates reflecting this robustness, with smooth and robustly estimated density curves.

In the left-bottom panel, the plots showcase the hybrid triweight kernels for all the families, juxtaposed with the classical triweight kernel plot. Similar to biweight kernels, the triweight kernel offers robust density estimates, generating smoother curves compared to the Epanechnikov kernel, often featuring a flatter peak.

Finally, in the right-bottom panel, the plots display the hy-

brid quadriweight kernels for all the families considered, including the classical quadriweight kernel plot. The quadriweight kernel assigns even more weight to central data points, resulting in exceptionally smooth density estimates. Consequently, the density curves may appear extremely smooth, potentially overlooking finer details in the data.

Like Figure 2, Figure 3 presents plots of kernel density estimations for different families of kernels, including classical kernels, using the COVID-19 dataset. Each subfigure focusses on a specific kernel family: Epanechnikov, biweight, triweight,

Table 7. Comparison of average global error of kernels in the cluster of families of hybrid kernels with the kernels in the family of classical kernels for four sample sizes of skewed distribution.

		AMISE*			
<i>n</i>	Family (<i>i</i>)	Epanechnikov	Biweight	Triweight	Quadriweight
10	Classical	0.00277444	0.00323586	0.00363370	0.00398071
	F1	0.00270974	0.00319246	0.00360074	0.003954
	F2	0.00264061	0.00314048	0.00355932	0.00391925
	F3	0.00257768	0.00308843	0.00351680	0.00388306
	F4	0.00252247	0.00303757	0.00347432	0.00384648
	F5	0.00247557	0.00298840	0.00343234	0.00380996
	F6	0.00243734	0.00294119	0.00339111	0.00377372
	F7	0.00240796	0.00289609	0.00335076	0.00373790
	F8	0.00238758	0.00285322	0.00331139	0.00370258
	F9	0.00237631	0.00281265	0.00327308	0.00366783
25	Classical	0.00088508	0.00103168	0.00115806	0.00126827
	F1	0.00086455	0.00101791	0.00114760	0.00125980
	F2	0.00084259	0.00100140	0.00113445	0.00124877
	F3	0.00082260	0.00098487	0.00112095	0.00123728
	F4	0.00080506	0.00096871	0.00110746	0.00122566
	F5	0.00079017	0.00095309	0.00109412	0.00121406
	F6	0.00077804	0.00093808	0.00108102	0.00120255
	F7	0.00076872	0.00092375	0.00106820	0.00119117
	F8	0.00076228	0.00091012	0.00105569	0.00117995
	F9	0.00075873	0.00089723	0.00104351	0.00116891
300	Classical	0.00008166	0.00009500	0.00010649	0.00011651
	F1	0.00007980	0.00009375	0.00010555	0.00011575
	F2	0.00007780	0.00009225	0.00010435	0.00011474
	F3	0.00007598	0.00009075	0.00010313	0.00011370
	F4	0.00007439	0.00008928	0.00010190	0.00011265
	F5	0.00007304	0.00008785	0.00010069	0.00011159
	F6	0.00007194	0.00008649	0.00009949	0.00011054
	F7	0.00007110	0.00008518	0.00009833	0.00010951
	F8	0.00007052	0.00008394	0.00009719	0.00010849
	F9	0.00007021	0.00008277	0.00009608	0.00010748
1500	Classical	0.00001663	0.00001931	0.00002162	0.00002364
	F1	0.00001625	0.00001906	0.00002143	0.00002348
	F2	0.00001585	0.00001876	0.00002119	0.00002328
	F3	0.00001549	0.00001846	0.00002095	0.00002307
	F4	0.00001516	0.00001816	0.00002070	0.00002286
	F5	0.00001489	0.00001787	0.00002045	0.00002265
	F6	0.00001467	0.00001760	0.00002021	0.00002244
	F7	0.00001450	0.00001734	0.00001998	0.00002223
	F8	0.00001439	0.00001709	0.00001975	0.00002202
	F9	0.00001433	0.00001685	0.00001953	0.00002182

and quadriweight.

In the top-left panel, the Epanechnikov kernel density estimations for all families, including the classical Epanechnikov kernel, are displayed. The Epanechnikov kernel is known for its triangular shape with a flattened peak, resulting in smooth density curves. The plots show variations among different families, indicating the impact of hybridisation on the density estimation process.

The top-right panel presents the biweight kernel density estimations for all families, including the classical biweight ker-

nel. The biweight kernel is recognised for its robustness to outliers, reflected in the smooth and robust density curves across different families. Again, variations among families are observed, showcasing the influence of hybridisation on density estimation.

Also, in the left-bottom panel, the triweight kernel density estimations for all families, including the classical triweight kernel, are depicted. Similar to the biweight kernel, the triweight kernel offers robust density estimates with smoother curves compared to the Epanechnikov kernel. Variations among

Table 8. Comparison of average global error of kernels in the cluster of families of hybrid kernels with the kernels in the family of classical kernels for four sample sizes of heavy-tailed distribution.

		AMISE*			
<i>n</i>	Family (<i>i</i>)	Epanechnikov	Biweight	Triweight	Quadriweight
10	Classical	0.00345962	0.00403499	0.00453108	0.00496379
	F1	0.00337894	0.00398087	0.00448998	0.00493048
	F2	0.00329273	0.00391605	0.00443834	0.00488715
	F3	0.00321427	0.00385115	0.00438531	0.00484202
	F4	0.00314542	0.00378773	0.00433234	0.00479641
	F5	0.00308694	0.00372642	0.00428000	0.00475087
	F6	0.00303926	0.00366755	0.00422858	0.00470568
	F7	0.00300263	0.00361131	0.00417826	0.00466101
	F8	0.00297722	0.00355786	0.00412917	0.00461697
	F9	0.00296317	0.00350727	0.00408140	0.00457364
25	Classical	0.00117845	0.00137364	0.00154191	0.00168865
	F1	0.00115111	0.00135530	0.00152799	0.00167737
	F2	0.00112187	0.00133333	0.00151048	0.00166268
	F3	0.00109526	0.00131131	0.00149250	0.00164738
	F4	0.00107191	0.00128980	0.00147453	0.00163192
	F5	0.00105208	0.00126899	0.00145678	0.00161647
	F6	0.00103592	0.00124902	0.00143933	0.00160114
	F7	0.00102352	0.00122993	0.00142226	0.00158599
	F8	0.00101494	0.00121179	0.00140560	0.00157105
	F9	0.00101022	0.00119462	0.00138939	0.00155635
300	Classical	0.00002810	0.00003269	0.00003665	0.00004010
	F1	0.00002746	0.00003226	0.00003632	0.00003983
	F2	0.00002677	0.00003175	0.00003591	0.00003949
	F3	0.00002615	0.00003123	0.00003549	0.00003913
	F4	0.00002560	0.00003072	0.00003507	0.00003877
	F5	0.00002514	0.00003023	0.00003465	0.00003840
	F6	0.00002476	0.00002976	0.00003424	0.00003804
	F7	0.00002447	0.00002932	0.00003384	0.00003769
	F8	0.00002427	0.00002889	0.00003345	0.00003734
	F9	0.00002416	0.00002848	0.00003306	0.00003699
1500	Classical	0.00000265	0.00000308	0.00000345	0.00000377
	F1	0.00000259	0.00000304	0.00000342	0.00000375
	F2	0.00000253	0.00000299	0.00000338	0.00000372
	F3	0.00000247	0.00000294	0.00000334	0.00000368
	F4	0.00000242	0.00000290	0.00000330	0.00000365
	F5	0.00000238	0.00000285	0.00000326	0.00000361
	F6	0.00000234	0.00000281	0.00000323	0.00000358
	F7	0.00000231	0.00000277	0.00000319	0.00000355
	F8	0.00000230	0.00000273	0.00000315	0.00000351
	F9	0.00000229	0.00000269	0.00000312	0.00000348

families are evident, illustrating the impact of hybridisation on density estimation.

Lastly, the right-bottom panel shows the quadriweight kernel density estimations for all families, including the classical quadriweight kernel. The quadriweight kernel assigns more weight to central data points, resulting in exceptionally smooth density estimates. Variations among families highlight the influence of hybridisation on the density estimation process.

Overall, the figure provides a comprehensive visualisation of the density estimation performance of different kernel fam-

ilies, demonstrating the impact of hybridisation on enhancing density estimation accuracy.

In Table 9, the average global error of hybrid kernels versus classical kernels is compared for the Life Span of Batteries dataset with $n = 40$. Hybrid kernels consistently demonstrate lower error values across all families and kernel types. As the index i increases, indicating different hybrid kernel configurations, there is a slight decrease in the average global error, indicating improved accuracy. These results highlight the potential of hybrid kernels for enhancing density estimation tasks in

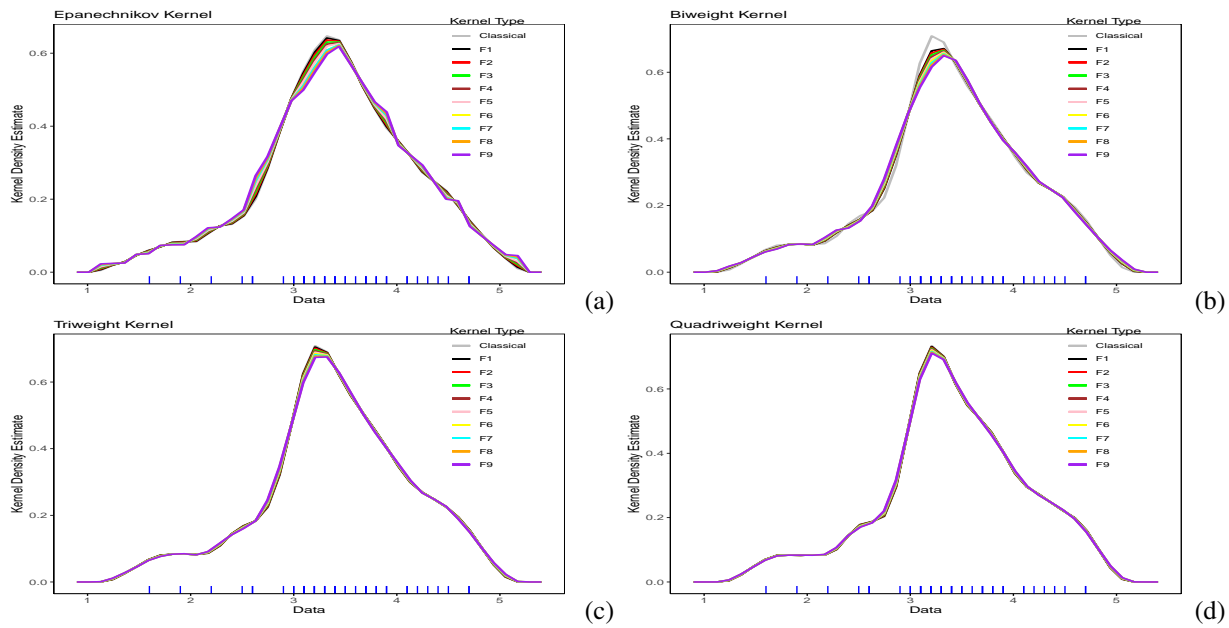


Figure 2. Plots of (a) Epanechnikov kernel for all the families including the classical Epanechnikov kernel, (b) biweight kernel for all the families including the classical biweight kernel, (c) triweight kernel for all the families including the classical triweight kernel, and (d) quadriweight kernel for all the families including the classical quadriweight kernel using the lifespan of car batteries in years [24].

Table 9. Comparison of average global error of kernels in the cluster of families of hybrid kernels with the kernels in the family of classical kernels for Life Span of Batteries dataset $n = 40$.

Family (i)	Epanechnikov	Biweight	Triweight	Quadriweight
Classical	0.0014719	0.0017152	0.0019248	0.0021077
F1	0.0014378	0.0016923	0.0019075	0.0020936
F2	0.0014014	0.0016649	0.0018857	0.0020753
F3	0.0013683	0.0016375	0.0018633	0.0020562
F4	0.0013392	0.0016107	0.0018409	0.0020370
F5	0.0013145	0.0015848	0.0018189	0.0020177
F6	0.0012943	0.0015599	0.0017970	0.0019986
F7	0.0012789	0.0015361	0.0017757	0.0019798
F8	0.0012682	0.0015135	0.0017550	0.0019611
F9	0.0012624	0.0014921	0.0017348	0.0019428

datasets like the Life Span of Batteries dataset, outperforming classical kernels.

Like Table 9, Table 10 shows the performance of each hybrid kernel family based on their average global error. Lower error values denote better prediction accuracy. For example, the Epanechnikov kernel shows that the classical family has the highest error, followed by F1, F2, and so forth, indicating decreasing error values across families. This suggests that the classical family performs the poorest, while F9 performs the best. Additionally, within each family, we can compare the performance of different kernel types. For instance, in the classical family, the Epanechnikov kernel exhibits the lowest error, suggesting relatively better performance compared to other kernels in the same family.

5. Discussion of results

In the domain of statistics, particularly within KDE, the accuracy of an estimator or kernel is a critical metric, often assessed through its error, commonly quantified as AMISE. The guiding principle is straightforward: a kernel with lower AMISE values exhibits greater accuracy compared to those with higher values. Our analysis, spanning Tables 5 through 10, is rooted in this principle, providing a meticulous comparative examination of various families of kernels within the cluster.

The detailed examination of kernel performance across various distributions and sample sizes via a Monte Carlo study, as outlined in Tables 5 through 8, provides compelling evidence of the superior accuracy of hybrid kernels over classical kernels. This analysis encompasses unimodal, multimodal, skewed, and heavy-tailed distributions, offering a comprehensive view of kernel efficiency through AMISE.

Starting with unimodal distributions, Table 5 illustrates that at smaller sample sizes, such as $n = 10$, classical kernels, in-

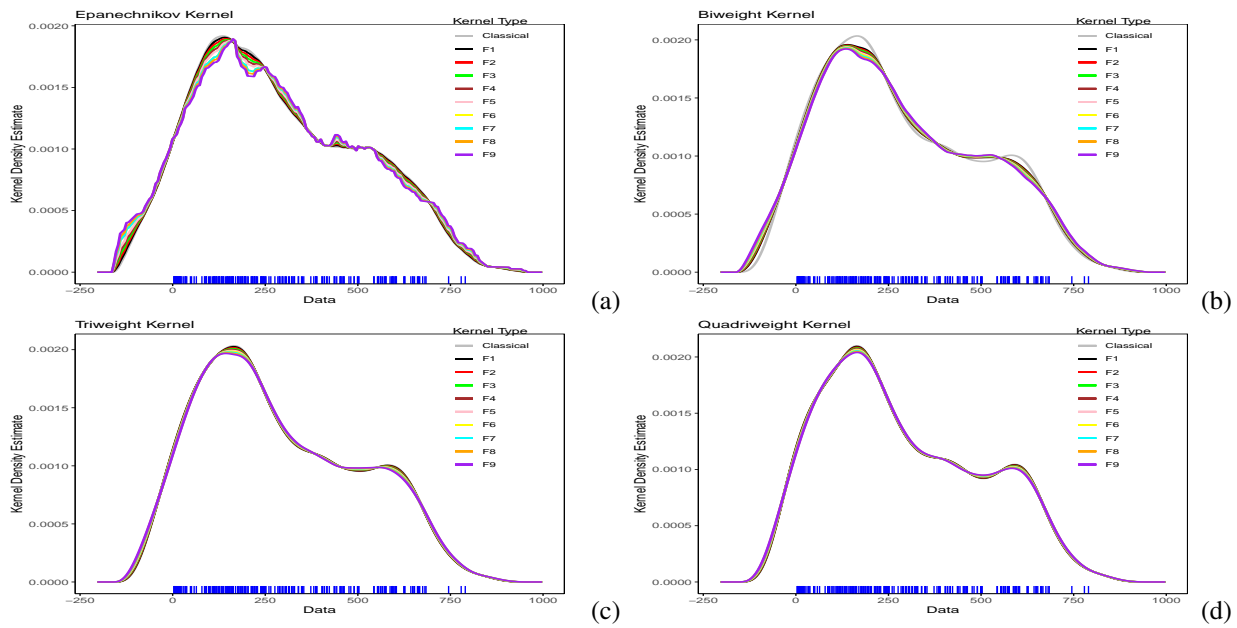


Figure 3. Plots of (a) Epanechnikov kernel for all the families including the classical Epanechnikov kernel, (b) biweight kernel for all the families including the classical biweight kernel, (c) triweight kernel for all the families including the classical triweight kernel, and (d) quadriweight kernel for all the families including the classical quadriweight kernel using the COVID-19 dataset [25].

Table 10. Comparison of average global error of kernels in the cluster of families of hybrid kernels with the kernels in the family of classical kernels for COVID-19 dataset $n = 214$.

Family (i)	AMISE*			
	Epanechnikov	Biweight	Triweight	Quadriweight
Classical	0.000000945384	0.000001100160	0.00000123352	0.00000134978
F1	0.000000923751	0.000001085650	0.00000122252	0.00000134087
F2	0.000000900594	0.000001068250	0.00000120866	0.00000132925
F3	0.000000879507	0.000001050800	0.00000119442	0.00000131713
F4	0.000000861009	0.000001033740	0.00000118019	0.00000130488
F5	0.000000845316	0.000001017240	0.00000116611	0.00000129265
F6	0.000000832546	0.000001001390	0.00000115228	0.00000128050
F7	0.000000822777	0.000000986248	0.00000113874	0.00000126849
F8	0.000000816062	0.000000971848	0.00000112552	0.00000125664
F9	0.000000812439	0.000000958220	0.00000111265	0.00000124498

cluding Epanechnikov, biweight, triweight, and quadriweight, exhibit higher average global errors compared to hybrid kernels. This trend persists as the sample size increases to $n = 25$, where hybrid kernels consistently outperform their classical counterparts. Among the hybrid kernels, family F9 emerges as the most effective, achieving the lowest average global error across all kernel types. This suggests that hybrid kernels are particularly adept at reducing estimation error for unimodal distributions across all sample scenarios. As the sample size grows to $n = 300$, the average global error decreases for all kernels, but hybrid kernels continue to show a notable advantage, albeit with a reduced margin. At the largest sample size of $n = 1500$, the average global error across all kernels becomes minimal, with hybrid kernels still providing slightly better performance, particularly for F9.

Moving to multimodal distributions, Table 6 reinforces the trend observed with unimodal distributions. Hybrid kernels

demonstrate superior performance at sample sizes $n = 10$, $n = 25$, $n = 300$, and $n = 1500$. At smaller sample sizes, hybrid kernels reduce the average global error compared to classical kernels, and this advantage continues as the sample size increases. The efficacy of hybrid kernels in managing complex multimodal distributions is evident, although the differences between classical and hybrid kernels become less pronounced as the sample size reaches $n = 1500$.

Table 7 further emphasises the superiority of hybrid kernels in the setting of skewed distributions. When compared to classical kernels, hybrid kernels considerably lower the average global error at small sample sizes, such as $n = 10$, demonstrating their efficacy in handling asymmetric data. As the sample size increases to $n = 25$, this benefit continues, with F9 exhibiting the smallest average global error. Hybrid kernels continue to be superior even as the sample size approaches $n = 300$, especially those belonging to families F6 to F9. Even with the

highest sample size of $n = 1500$, hybrid kernels still outperform classical kernels by a small margin, even when the average global error for all kernels approaches zero.

Table 8 extends this analysis to heavy-tailed distributions, where hybrid kernels once again prove superior. The improvement in accuracy with more intricate hybrid kernel configurations underscores their efficacy in heavy-tailed density estimation tasks.

The examination of real-life datasets, such as the lifespan of car batteries with $n = 40$ and COVID-19 data with $n = 214$, corroborates the superior performance of hybrid kernels. For the car battery lifespan dataset, as given in Table 7, hybrid kernels consistently outperform classical kernels across all types, with F9 achieving the lowest average global error. A similar trend is observed with the COVID-19 dataset given in Table 8, where hybrid kernels demonstrate lower average global errors compared to classical kernels, with F9 again providing the best performance.

In essence, the comparative results across various distributions and real-life datasets highlight the superior performance of hybrid kernels. These kernels consistently achieve lower average global errors, especially with increasing sample sizes, emphasising their effectiveness in capturing complex underlying distributions. The findings reinforce the advantages of hybrid kernels in improving estimation accuracy across different scenarios, with their performance benefits becoming more pronounced as sample sizes grow.

6. Conclusion

This study introduced a novel family of hybrid beta polynomial kernels, which consistently outperformed traditional polynomial kernels in density estimation tasks. Through a comprehensive analysis of both simulated and real-life data, these hybrid kernels demonstrated superior accuracy, particularly in reducing average global errors across unimodal, multimodal, skewed, and heavy-tailed distributions. The findings highlight the robustness and versatility of hybrid kernels, showing their effectiveness in a wide range of scenarios, especially as sample sizes increase. Their ability to handle complex distributions and outperform classical kernels makes them a promising advancement in density estimation methodologies. The study's key contribution lies in its systematic comparison of hybrid and classical kernels, with a specific focus on heavy-tailed distributions. The results emphasise the superior performance of hybrid kernels, encouraging their adoption in density estimation and providing a solid foundation for future research in areas such as high-dimensional data analysis and advanced estimation techniques. Further studies could also investigate the potential of combining these kernels with other advanced estimation techniques to enhance their performance across even broader scenarios.

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] 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. 238–247. <https://sci-hub.se/10.2307/1403797>.
- [2] H. Akaike, "An approximation to the density functions", Annals of the Institute of Statistical Mathematics 6 (1954) 127. <https://doi.org/10.1007/BF02900741>.
- [3] 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>.
- [4] 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>.
- [5] B. W. Silverman, "Density Estimation for Statistics and Data Analysis", Biometrical Journal 30 (1986) 876. <https://doi.org/10.1002/bimj.4710300745>.
- [6] M. Jiang & S. P. Provost, "A hybrid bandwidth selection methodology for kernel density estimation", Journal of Statistical Computation and Simulation 84 (2014) 614. <https://doi.org/10.1080/00949655.2012.721366>.
- [7] I. S. Abramson, "On bandwidth variation in kernel estimates—a square root law", Annals of Mathematical Statistics 27 (1982) 1217. <https://doi.org/10.1214/aos/1176345986>.
- [8] C. J. Stone, "An asymptotically optimal window selection rule for kernel estimates", Annals of Statistics 12 (1984) 1285. <https://doi.org/10.1214/aos/1176346792>.
- [9] J. S. Marron & W. J. Padgett, "Asymptotically optimal bandwidth selection for kernel density estimators from randomly-censored samples", Annals of Statistics 15 (1987) 1520. <https://doi.org/10.1214/aos/1176350607>.
- [10] M. P. Wand & M. C. Jones, *Kernel Smoothing*, Chapman & Hall, London, UK, 1994, pp. 1–224. <https://doi.org/10.1201/b14876>.
- [11] J. Simonoff, *Smoothing Methods in Statistics*, Springer, New York, USA, 1996, pp. 1–339. <https://link.springer.com/book/10.1007/978-1-4612-4026-6>.
- [12] A. W. Bowman & A. Azzalini, *Applied Smoothing Techniques for Data Analysis: The Kernels Approach with S-Plus Illustration*, Oxford University Press, UK, 1997, pp. 1–168. <https://doi.org/10.1093/oso/9780198523963.001.0001>.
- [13] 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>.
- [14] I. U. Siloko, C. C. Ishiekwene & F. O. Oyegue, "New gradient methods for bandwidth selection in bivariate kernel density estimation", Mathematics and Statistics 6 (2018) 1. <https://www.hrpub.org/download/20171230/MS1-13409787.pdf>.
- [15] 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>.
- [16] B. A. Afere & F. O. Oyegue, "On the construction of the family of d -dimensional spherically symmetric polynomial kernels", Palestine Journal of Mathematics 8 (2019) 286. <https://pjm.ppu.edu/sites/default/files/papers/24%20%20%20Adel%20ICMS15b.pdf>.
- [17] I. U. Siloko, W. Nwankwo & E. A. Siloko, "A new family of kernels from the beta polynomial kernels with applications in density estimation", International Journal of Advances in Intelligent Informatics 6 (2010) 235. <https://www.semanticscholar.org/reader/8d742f2fc7aab2c8ee062f672418231dbd4475f7>.
- [18] J. Lin & X. Wu, "A hybrid nonparametric multivariate density estimator with applications to risk management", Econometrics Reviews 42 (2024) 301. <https://doi.org/10.1080/07474938.2024.2334119>.
- [19] B. A. Afere, "A new family of hybrid classical polynomial kernels in density estimation", International Journal of Research and Innovation Applied Sciences VI (2021) 145. <https://rsisinternational.org/journals/ijrias/DigitalLibrary/Vol.6&Issue1/145-151.pdf>.
- [20] B. A. E. Afere, "On the fourth-order hybrid beta polynomial kernels in kernel density estimation", Journal of the Nigerian Society of Physical Sciences 6 (2024) 1631. <https://doi.org/10.46481/jnsps.2024.1631>.

- [21] D. W. Scott, *Multivariate Density Estimation: Theory, practice and visualization*, John Wiley & Sons Inc., New York, USA, 1992, pp. 1–317. <https://onlinelibrary.wiley.com/doi/book/10.1002/9780470316849>.
- [22] J. S. Marron & M. P. Wand, “Exact mean integrated squared error”, *Annals of Statistics* **20** (1992) 712. <https://doi.org/10.1214/aos/1176348653>.
- [23] A. J. Izenman, *Modern Multivariate Statistical Techniques: regression, Classification, and Manifold Learning*, Springer, New York, USA, 2008, pp. 1–733. <https://link.springer.com/book/10.1007/978-0-387-78189-1>.
- [24] C. C. Ishiekwene & B. A. E. Afere, “Higher-order window width selectors for empirical data”, *Journal of Nigerian Statistical Association* **14** (2001) 69. <https://scholar.google.com/citations?user=I3DojTQAAAAJ>.
- [25] P. Osatohanmwun, E. Efe-Eyefia, F. O. Oyegun, J. E. Osemwenkhae, S. M. Ogbonmwun & B. A. Afere, “The exponentiated gumbel–weibull Logistic distribution with application to Nigeria’s COVID–19 infections data”, *Annals of Data Science* **9** (2022) 909. <https://doi.org/10.1007/s40745-022-00373-0>.