Nonparametric Statistics
Vol. 17, No. 5, July 2005, 555–582

Taylor & Francis
Taylor & Francis Group

Convergence rates for uniform B-spline density estimators on bounded and semi-infinite domains

RIAD MASRI† and RICHARD A. REDNER*‡

†Department of Mathematics, The University of Texas at Austin, Austin, TX 78722, USA
‡Department of Mathematical and Computer Sciences, The University of Tulsa, Tulsa, OK 74104-3189

(Received October 2002; revised October 2003; in final form July 2004)

Convergence rates for B-spline nonparametric density estimators on bounded and semi-infinite domains are discussed. We show how B-spline estimators can be adjusted to account for edge effects and then determine the mean integrated squared error for the adjusted estimator and its derivatives.

Keywords: Asymptotic mean integrated squared error; Nonparametric density estimation; Splines

AMS 1991 Subject Classifications: 62G07, 68U05

1. Introduction

Simulations that generate very large data sets in one and many dimensions are increasingly common. Nonparametric density estimates based on these data sets are often required and estimators that can be generated and manipulated efficiently are needed. In computer graphics, for example, nonparametric density estimates over surfaces can be used to represent lighting functions [1]. These simulations may generate as many as 50–100,000 data points for each light in the scene. Another application area is that of statistical genetics where simulations are needed for hypothesis tests when the exact distribution of the test statistics is not known [2]. For example, the likelihood ratio test statistics is, under reasonable conditions, asymptotically chi-squared. But in the case that the parameter for the null hypothesis is on the boundary of the parameter domain, the distribution of the test statistics may be unknown.

The fact that spline functions can be efficiently evaluated on a digital computer has led to the use of splines for many statistical applications [3–9] and for density estimation [10]. We note that the paper by Lii [7] uses spline interpolation of the cumulative distribution function and shows that the bias is $O(h^3)$. But this result depends on knowing certain endpoint information. If this information is incorrect, then the estimate would have much larger bias at the endpoints.

In addition to these papers, a B-spline nonparametric density estimator with uniformly spaced knots convenient for large data sets was discussed by Gehringer and Redner [11]. These

*Corresponding author. Tel.: +1-918-631-2986; Email: rredner@utulsa.edu

Nonparametric Statistics
ISSN 1048-5252 print/ISSN 1029-0311 online
http://www.tandf.co.uk/journals
DOI: 10.1080/10485250500039239

ideas were later extended to density function estimation on metric spaces using partitions of unity in Redner and Gehringer [11]. The asymptotic rates of convergence for these schemes in one and multiple dimensions have also been investigated by Redner [see refs. 12, 13]. The results in these papers establish the fact that the estimators converge rapidly, and furthermore give approximately the same rate of convergence as kernel density estimators in one and many dimensions. Yet the final B-spline density estimate depends only on a relatively small number of statistics. Additional details of the computational advantage of B-spline estimators can be found in ref. [12].

As is the case with kernel density estimates, the affects of boundaries must be accounted for. Note however, that by definition, the B-spline density estimates naturally have support within the interval specified and so, for example, the reflection techniques as presented by Schuster [14] are not applicable. But other ideas for correcting for edge effects for kernel density estimators, which include a transformation approach by Marron and Ruppert [15], the pseudodata method of Cowling and Hall [16] and a Hilbert space projection method of Djojosugito and Speckman [17], could be adapted to the B-spline boundary problem. Additional ideas include the beta kernel estimator presented by Chen [18] and the boundary kernel method of Müller [19] could also be used. The reader is referred to the paper Cheng *et al.* [20] for several additional references.

In this paper, however, we do not pursue any of these methods since it seem unlikely that any of these methods would appropriately correct the boundary effects on the derivatives of the density estimate. Instead, we introduce a boundary correction method specifically designed for the B-spline density estimate on finite and semi-infinite intervals. We also establish the asymptotic rate of convergence of the estimator and all of its continuous derivatives.

In section 2, we introduce the corrected estimator and show how the correction coefficients can be uniquely determined. Our main results on the asymptotic integrated squared bias and variance for semi-infinite and bounded intervals are presented in sections 3 and 4. The proof of the results for the semi-infinite case of section 3 are given in sections 5, 6 and 7 and is followed by a brief discussion in section 8.

2. The B-spline density estimator with boundary correction

Let $N^m(x)$ be the mth order normalized uniform B-spline associated with the evenly spaced knots $0, 1, 2, \ldots, m$, *i.e.*, for each real number x,

$$N^m(x) = \sum_{i=0}^{m} (-1)^i \binom{m}{i} \frac{(x-i)_+^{m-1}}{(m-1)!},$$

where

$$(z)_+ = \begin{cases} z, & \text{if } z < 0; \\ 0, & \text{otherwise.} \end{cases}$$

We note that $\int_{-\infty}^{\infty} N^m(x)\,\mathrm{d}x = 1$, $N^m(x) \geq 0$ and that $N^m(x)$ is $m-2$ times continuously differentiable [21, 22]. In this application, we will need scaled and shifted copies of the basis functions. So let h be a positive real number, $x_i = (i-m)h$ for each integer i and define

$$B_i^m(x) = N^m\left(\frac{x - x_i}{h}\right) = N^m\left(\frac{x}{h} + m - i\right).$$

Then for $m \geq 1$, $0 \leq B_i^m(x) \leq 1$ and $\sum_{i=-\infty}^{\infty} B_i^m(x) = 1$ for each x. When $m \geq 2$, the basis functions are continuous and hence form a partition of unity.

Therefore, let $X_1, X_2, \ldots, X_N$ and X be independent identically distributed random variables from a continuous probability density function f on a closed interval I. Define an estimate $\hat{f}_N$ of f as

$$\hat{f}_N(x) = \sum_i \frac{\alpha_i}{b_i} B_i^m(x),$$

where b_i and α_i are defined by

$$b_i = \int_I B_i^m(x)\,\mathrm{d}x \quad \text{and} \quad \alpha_i = \frac{1}{N}\sum_{k=1}^{N} B_i^m(X_k)$$

for each integer i and where $\sum_i$ is the sum over all indices for which $b_i \neq 0$. Note that $\hat{f}_N(x)$ is a probability density function and under the natural assumptions in ref. [1], this estimator converges in mean integrated squared error to the true underlying density function. For $I = (-\infty, \infty)$, the asymptotic mean integrated square error (AMISE) was determined in ref. [12] for $\hat{f}_N$ and all of its nontrivial derivatives. But these results do not apply to the case that I is a bounded or semi-infinite interval and in fact the boundary has a deleterious effect on the size of the AMISE. The goal of this paper is to introduce a modified estimator $\hat{f}_C(x)$ and to determine the AMISE for bounded and semi-infinite intervals.

To reach this goal, correction terms will be introduced, so let x_0 be a real number, let k be a whole number and define

$$a_{i,k}(x_0) \equiv \int_{-\infty}^{\infty} \frac{(x-x_0)^k}{k!}\,\frac{B_i^m(x)}{h}\,\mathrm{d}x$$

and

$$\alpha_{i,k}(x_0) \equiv \int_0^{\infty} \frac{(x-x_0)^k}{k!} B_i^m(x)\,\mathrm{d}x.$$

We will also use $a_{i,k} = a_{i,k}(0)$ and $\alpha_{i,k} = \alpha_{i,k}(0)$. Note also that $\alpha_{i,k}(x_0)/b_i = a_{i,k}(x_0)$ for $i \geq m$.

The proof of the following theorem can be found in Appendix A.

THEOREM 1 *If $m \geq 2$, then the system of equations*

$$\sum_{j=i}^{m+i-1} c_{i,j+1-i}\,\alpha_{j,k} = a_{i,k} b_i - \alpha_{i,k} \quad \text{for } i = 1, \ldots, m-1 \text{ and } k = 0, \ldots, m-1 \tag{1}$$

has a unique solution.

Given the $m(m-1)$ unknown constants $\{\{c_{ij}\}_{i=1}^{m-1}\}_{j=1}^{m}$ defined by solving the system of equations in Theorem 1, the new B-spline density estimator can now be defined. Let $I = [0, \infty)$

and define

$$B_i^m(x) = N^m\left(\frac{x}{h} + m - i\right) \quad \text{and} \quad b_i = \int_0^\infty B_i^m(x)\,\mathrm{d}x$$

for $i \geq 1$. We define

$$\hat{f}_N^*(x) = \sum_{i=1}^{m-1} \alpha_{L_i}^* \frac{B_i^m(x)}{b_i} + \sum_{i=1}^{\infty} \alpha_i \frac{B_i^m(x)}{b_i} \tag{2}$$

and

$$\hat{f}_C(x) = \hat{f}_N^*(x) - \bar{\alpha}\hat{f}_N(x), \tag{3}$$

where

$$\alpha_{L_i}^* = \sum_{j=i}^{m+i-1} c_{i,j+1-i}\alpha_j$$

for $i = 1, \ldots, m-1$ and

$$\bar{\alpha} = \sum_{i=1}^{m-1} \alpha_{L_i}^*.$$

The first term in the expression of $\hat{f}_N^*(x)$ in equation 2 is added to remove the effects of the left boundary and the last term of equation (3) is added so that $\hat{f}_C(x)$ integrates to 1.

We have used Mathematica [23] to determine the correction terms defined by equation (1). When $m = 2$, the correction term $\sum_{i=1}^{m-1} \alpha_{L_i}^* B_i^m(x)/b_i$ is

$$\alpha_{L_1}^* \frac{B_1^2(x)}{b_1},$$

where

$$\alpha_{L_1}^* \frac{1}{2}\alpha_1 - \frac{1}{4}\alpha_2.$$

In the case that $m = 3$, the correction term is

$$\sum_{i=1}^{2} \alpha_{L_i}^* \frac{B_i^3(x)}{b_i},$$

where

$$\alpha_{L_1}^* = \frac{89}{18}\alpha_1 - \frac{23}{18}\alpha_2 + \frac{13}{54}\alpha_3$$

and

$$\alpha_{L_2}^* = \frac{13}{37}\alpha_2 - \frac{5}{12}\alpha_3 + \frac{55}{444}\alpha_4.$$

The results for splines of order 4, 5 and 6 are included in Appendix B.

3. Assumptions and main results

Throughout this paper, unless otherwise noted, we make the following assumptions

ASSUMPTION 1 *Both m and n are whole numbers with $m \geq n+2$.*

ASSUMPTION 2 *Given an interval I, f is $n+2$ times continuously differentiable on I, $f^{(n+2)}$ is absolutely continuous and $f^{(k)} \in L_1(I)$ and $f^{(k)} \in L_2(I)$ for $0 \leq k \leq n+3$.*

Except in section 4, we will assume that $I = [0, \infty)$. Then the following two theorems represent the main results of this paper. Technical results needed to prove these facts are presented in section 5 and the proofs of Theorems 2 and 3 are given in sections 6 and 7, respectively.

THEOREM 2 *If $m \geq n+3$, the integrated squared bias is*

$$\int_0^\infty (E(\hat{f}_C^{(n)}(x)) - f^{(n)}(x))^2 \mathrm{d}x$$
$$= h^4 \left(\frac{m}{12}\right)^2 \int_0^\infty (f^{(n+2)}(x) + f'(0) f^{(n)}(x))^2 \mathrm{d}x + O(h^5) + O\left(\frac{h^2}{Nh^n}\right)$$

and if $m = n+2$, the integrated squared bias is

$$\int_0^\infty (E(\hat{f}_C^{(n)}(x)) - f^{(n)}(x))^2 \mathrm{d}x = \frac{h^4}{720} \int_0^\infty (f^{(n+2)}(x))^2 \mathrm{d}x$$
$$+ h^4 \left(\frac{m}{12}\right)^2 \int_0^\infty (f^{(n+2)}(x) + f'(0) f^{(n)}(x))^2 \mathrm{d}x + O(h^5) + O\left(\frac{h^2}{Nh^n}\right).$$

We note that in the unbounded case [see ref. 12], the integrated variance of the estimator $\hat{f}_N^{(n)}$ was shown to be

$$\frac{D_{m,n}}{Nh^{1+2n}} + O\left(\frac{1}{Nh^{2n}}\right).$$

We have similar results for the estimator $\hat{f}_C^{(n)}$.

THEOREM 3 *The integrated variance of the estimator $\hat{f}_C^{(n)}$ is*

$$\frac{D_{m,n}}{Nh^{1+2n}} + O\left(\frac{\sqrt{h}}{Nh^{1+2n}}\right),$$

where $D_{m,n}$ is independent of N and h and examples of values of $D_{m,n}$ can be found in table B4 in Appendix B.

As an immediate corollary we have the following result.

COROLLARY 1 *If $h \to 0$ and $Nh^{1+2n} \to \infty$ as $N \to \infty$, then $\hat{f}_C^{(n)}$ converges to $f^{(n)}$ in mean integrated squared error.*

4. Auxiliary results

In this section only we define $I = [a, b]$ and assume that Assumption 2 holds for $I = [a, b]$. Then the ideas and results of section 2 can be directly extended. In this case, we need to use values of h that allow knots to fall at both a and b. So let K_1 be a positive integer and define $h = (b - a)/K_1$, $B_i^m(x) = N(x/h + m - ah - i)$ and $K = K_1 + m - 1$. Then the basis functions $\{B_i^m\}_{i=1}^K$ have support on $[a, b]$. So define

$$\hat{f}_N^*(x) = \sum_{i=1}^{m-1} \alpha_{L_i}^* \frac{B_i^m(x)}{b_i} + \sum_{i=1}^{K} \alpha_i \frac{B_i^m(x)}{b_i} + \sum_{i=K-m+2}^{K} \alpha_{R_i}^* \frac{B_i^m(x)}{b_i}$$

and

$$\hat{f}_C(x) = \hat{f}_N^*(x) - \bar{\alpha} \hat{f}_N(x),$$

where

$$\alpha_{L_i}^* = \sum_{j=i}^{m+i-1} c_{i,j+1-i}\alpha_j,$$

$$\alpha_{R_i}^* = \sum_{j=i-m+1}^{i} c_{K-i+1,i-j+1}\alpha_j,$$

and

$$\bar{\alpha} = \sum_{i=1}^{m-1} \alpha_{L_i}^* + \sum_{i=K-m+2}^{K} \alpha_{R_i}^*.$$

THEOREM 4 *If $m \geq n + 3$,*

$$\int_a^b \left(E\left(\hat{f}_C^{(n)}(x)\right) - f^{(n)}(x)\right)^2 \mathrm{d}x$$
$$= h^4 \left(\frac{m}{12}\right)^2 \int_a^b \left(f^{(n+2)}(x) + (f'(a) - f'(b)) f^{(n)}(x)\right)^2 \mathrm{d}x + O(h^5) + O\left(\frac{h^2}{Nh^n}\right)$$

and if $m = n + 2$,

$$\int_a^b \left(E\left(\hat{f}_C^{(n)}(x)\right) - f^{(n)}(x)\right)^2 \mathrm{d}x$$
$$= \frac{h^4}{720} \int_a^b \left(f^{n+2}(x)\right)^2 \mathrm{d}x + h^4 \left(\frac{m}{12}\right)^2$$
$$\times \int_a^b \left(f^{(n+2)}(x) + (f'(a) - f'(b)) f^{(n)}(x)\right)^2 \mathrm{d}x + O(h^5) + O\left(\frac{h^2}{Nh^n}\right).$$

THEOREM 5 *The integrated variance of the estimator $\hat{f}_C^{(n)}$ is*

$$\frac{D_{m,n}}{Nh^{1+2n}} + O\left(\frac{\sqrt{h}}{Nh^{1+2n}}\right),$$

where $D_{m,n}$ is independent of N and h and examples of values can be found in table B4 in Appendix B.

Convergence of the estimator and its derivatives is assured under the conditions of Corollary 1.

5. Preliminary theorems and lemmas

We now return to our analysis of the case that $I = [0, \infty)$. For each of the next two theorems, we define the sequence of points $t_i = hi$ for $i = 0, 1, 2, \dots$ and let ξ_i and η_i denote arbitrary points in the ith interval, *i.e.*, $t_{i-1} \le \xi_i \le t_i$ and $t_{i-1} \le \eta_i \le t_i$.

THEOREM 6 [12] *Suppose that the function ϕ is absolutely continuous on $[0, \infty)$ and that ϕ and ϕ' are in $L_1[0, \infty)$. Then the following sum exists and may be written as*

$$\sum_{i=1}^{\infty} \phi(\xi_i)h = \int_0^{\infty} \phi(x)\,\mathrm{d}x + O(h\|\phi'\|_1).$$

THEOREM 7 [12] *Let f and g satisfy f', g', fg, and $(fg)' \in L_1[0, \infty)$. Then*

$$\sum_{i=1}^{\infty} f(\xi_i)g(\eta_i)h = \int_0^{\infty} f(x)g(x)\,\mathrm{d}x + O(h\|(fg)'\|_1) + O(h\|f'\|_1(\inf|g| + \|g'\|_1)).$$

Please note that if the points ξ_i and η_i lie in overlapping intervals of the form $[t_{i-k}, t_i]$ for a fixed number k, then the conclusions of Theorems 6 and 7 still hold.

THEOREM 8 [12] *Let $\{a_i\}_{i=-\infty}^{\infty}$ be any bi-infinite sequence of real numbers. Then for $0 \le n \le m$, $\mathrm{d}^n/\mathrm{d}x^n \sum_{i=1}^{\infty} a_i B_i^m(x) = 1/h^n \sum_{i=1}^{\infty} (\Delta^n a_i) B_i^{m-n}(x)$, where $\mathrm{d}/\mathrm{d}x$ denotes the right-hand derivative and Δ denotes the forward difference operator.*

For $n \ge 0$, we now define $g_n(x, x_0) = \sum_{i=-\infty}^{\infty} a_{i,n}(x_0) B_i^m(x)$, where $B_i^m(x) = N^m(x/h + m - i)$. Then the following theorem is a straightforward generalization of the work by Redner in ref. [12] where the result is proven for $x_0 = 0$ and the basis functions are shifted by 1/2. Note also that all B-spline basis functions are differentiated using a right-handed derivative since our B-splines are defined to be right continuous.

THEOREM 9 *Let x_0 be a real number. For $m \ge n + 1$, $g_n^{(n)}(x, x_0) = 1$.*

If $m \ge n + 2$, $g_{n+1}^{(n)}(x, x_0) = x - x_0$.

If $m = n + 2$, $g_{n+2}^{(n)}(x, h/2) = h^2 \dfrac{2m+1}{24}$

and if $m \ge n + 3$,

$$g_{n+2}^{(n)}(x, x_0) = \frac{(x - x_0)^2}{2} + \frac{h^2 m}{12}.$$

We now develop several facts concerning properties of the estimate of the probability density function f and some of its components. Therefore, we observe that if f is M times continuously differentiable on $[0, \infty)$ and x_l is any real number, then by Taylor's theorem

$$\begin{aligned} E(\alpha_i) &= \int_0^\infty f(x) B_i^m(x)\,\mathrm{d}x \\ &= \int_0^\infty \left(\sum_{k=0}^{M-1} f^{(k)}(x_l) \frac{(x-x_l)^k}{k!} + f^{(M)}(\xi_x) \frac{(x-x_l)^M}{M!} \right) B_i^m(x)\,\mathrm{d}x \\ &= \sum_{k=0}^{M-1} f^{(k)}(x_l)\alpha_{i,k}(x_l) + \int_0^\infty f^{(M)}(\xi_x) \frac{(x-x_l)^M}{M!} B_i^m(x)\,\mathrm{d}x, \end{aligned}$$

where ξ_x is between x and x_l for each x. If M is even, then by the integral mean value theorem

$$E(\alpha_i) = \sum_{k=0}^{M-1} f^{(k)}(x_l)\alpha_{i,k}(x_l) + f^{(M)}(\xi_i)\alpha_{i,M}(x_l), \tag{4}$$

where ξ_i lies in the support of $B_i^m(x)$.

When k is even, define

$$p_i^k = \begin{cases} \displaystyle\sum_{j=i}^{m+i-1} c_{i,j+1-i} f^{(k)}(\xi_j)\alpha_{j,k}(x_l)/b_i + f^{(k)}(\xi_i)\alpha_{i,k}(x_l)/b_i & \text{if } 1 \le i \le m-1; \\ f^{(k)}(\xi_i)\alpha_{i,k}(x_l) & \text{otherwise.} \end{cases} \tag{5}$$

The case that k is odd can be handled in a similar fashion and for all $k \ge 0$, $p_i^k = O(h^k)$. The following lemma can now be established in a straightforward manner using equations (4) and (5) and their extensions for the odd case and by considering the cases that M is even or odd.

LEMMA 1 *Let f be M times continuously differentiable on* $[0, \infty)$. *Then for* $x_l \ge 0$,

$$E\left(\hat{f}_N^*(x)\right) = \sum_{k=0}^{M-1} f^{(k)}(x_l) g_k(x, x_l) + \sum_{i=1}^{\infty} p_i^M B_i^m(x).$$

LEMMA 2 *For* $x_l \ge 0$,

$$E\left(\hat{f}_N^{*(n)}(x)\right) = f^{(n)}(x_l) + f^{(n+1)}(x_l)(x - x_l) + \sum_{i=1}^{\infty} p_i^{n+2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} B_i^m(x).$$

Proof Since $g_n^{(n)}(x, x_l) = 1$ and $g_{n+1}^{(n)}(x, x_l) = x - x_l$, then the result follows by Lemma 1 with $M = n + 2$. ■

LEMMA 3

$$\int_0^\infty \left(E(\hat{f}_N^{(n)}(x))\right)^2 \mathrm{d}x = \int_0^\infty \left(f^{(n)}(x)\right)^2 \mathrm{d}x + O(h).$$

Proof Using the fact that $\hat{f}_N^*(x) = \hat{f}_N(x)$ on $x \ge (m-1)h$, the result follows in a straightforward way from ref. [12] ■

In addition to this theorem, we will need two technical results that will be used to determine the AMISE.

LEMMA 4 *Let* $\{\{c_{i,j}\}_{i=1}^{m-1}\}_{j=1}^{m}$ *satisfy equation* (1). *Then for* $i = 1, \ldots, m-1$, $\sum_{j=i}^{m+i-1} c_{i,j+1-i} b_j = 0$.

Proof

$$\sum_{j=i}^{m+i-1} c_{i,j+1-i} b_j = \sum_{j=i}^{m+i-1} c_{i,j+1-i} \alpha_{j,0} = a_{i,0} b_i - \alpha_{i,0} = 1 b_i - b_i = 0.$$

■

LEMMA 5 *Let* x_0 *be any real number and let* $\{\{c_{i,j}\}_{i=1}^{m-1}\}_{j=1}^{m}$ *satisfy equation* (1). *Then*

$$\sum_{j=i}^{m+i-1} \frac{c_{i,j+1-i} \alpha_{j,k}(x_0)}{b_i} + \frac{\alpha_{i,k}(x_0)}{b_i} = a_{i,k}(x_0)$$
$$\textit{for } i = 1, \ldots, m-1 \ \textit{ and } \ k = 0, \ldots, m-1. \tag{6}$$

Proof We first observe by an application of the binomial expansion theorem that

$$\alpha_{j,k}(x_0) = \sum_{p=0}^{k} \frac{(-x_0)^{k-p}}{(k-p)!} \alpha_{j,p}.$$

Then, since $\{\{c_{ij}\}_{i=1}^{m-1}\}_{j=1}^{m}$ satisfies equation (1), it follows for $i = 1, \ldots, m-1$ and $k = 0, \ldots, m-1$, that

$$\begin{aligned}\sum_{j=i}^{m+i-1} \frac{c_{i,j+1-i} \alpha_{j,k}(x_0)}{b_i} + \frac{\alpha_{i,k}(x_0)}{b_i} &= \sum_{p=0}^{k} \frac{(-x_0)^{k-p}}{(k-p)!} \left(\sum_{j=i}^{m+i-1} \frac{c_{i,j+1-i} \alpha_{j,p}}{b_i} + \frac{\alpha_{i,p}}{b_i} \right) \\ &= \sum_{p=0}^{k} \frac{(-x_0)^{k-p}}{(k-p)!} \int_{-\infty}^{\infty} \frac{x^p}{p!} \frac{B_i^m(x)}{h} \, \mathrm{d}x = a_{i,k}(x_0).\end{aligned}$$

■

In Lemma 6, we consider the expression $E(\bar{\alpha})$.

LEMMA 6

$$E(\bar{\alpha}) = -\frac{m}{12} h^2 f'(0) + O(h^3).$$

Proof Using Lemma 4 and Theorem C1 we see that

$$\begin{aligned}E(\bar{\alpha}) &= \sum_{i=1}^{m-1} \sum_{j=i}^{m+i-1} c_{i,j+1-i} E(\alpha_j) \\ &= \sum_{i=1}^{m-1} \sum_{j=i}^{m+i-1} c_{i,j+1-i} \left(f(0) b_j + f'(0) \alpha_{j,1} + O(h^3) \right) \\ &= -h^2 f'(0) \frac{m}{12} + O(h^3).\end{aligned}$$

■

LEMMA 7 *Let $X_1, \ldots, X_N$ and X be independent identically distributed random variables. Then*

$$E(\alpha_i\alpha_j) = \left(1 - \frac{1}{N}\right) E(\alpha_i)E(\alpha_j) + \frac{1}{N} E(B_i(X)B_j(X))$$

$$E(\alpha_i\alpha_j\alpha_k) = \frac{(N-1)(N-2)}{N^2} E(\alpha_i)E(\alpha_j)E(\alpha_k) + \frac{1}{N^2} E(B_i(X)B_j(X)B_k(X)) + \frac{(N-1)}{N^2} \big(E(B_i(X)B_j(X))E(B_k(X)) + E(B_i(X))E(B_j(X)B_k(X)) + E(B_i(X)B_k(X))E(B_j(X))\big),$$

and

$$\begin{aligned} E(\alpha_i\alpha_j\alpha_k\alpha_l) = {} & \frac{N(N-1)(N-2)(N-3)}{N^4} \big(E(\alpha_i)E(\alpha_j)E(\alpha_k)E(\alpha_l)\big) \\ & + \frac{N(N-1)(N-2)}{N^4} \big(E(B_iB_j)E(B_k)E(B_l) + E(B_iB_k)E(B_j)E(B_l) \\ & + E(B_iB_l)E(B_j)E(B_k) + E(B_jB_k)E(B_i)E(B_l) + E(B_jB_l)E(B_i)E(B_k) \\ & + E(B_kB_l)E(B_i)E(B_j)\big) + \frac{N(N-1)}{N^4} \big(E(B_iB_j)E(B_kB_l) \\ & + E(B_iB_k)E(B_jB_l) + E(B_iB_l)E(B_jB_k)\big) + \frac{N(N-1)}{N^4} \big(E(B_iB_jB_k)E(B_l) \\ & + E(B_iB_jB_l)E(B_k) + E(B_iB_kB_l)E(B_j) + E(B_jB_kB_l)E(B_i)\big) \\ & + \frac{N}{N^4} E(B_iB_jB_kB_l). \end{aligned}$$

Proof Recalling that $\alpha_i = 1/N \sum_{k=1}^{N} B_i(X_k)$, these results are easily obtained using the independence of X_k and X_l when $k \neq l$ and the application of simple counting arguments [see also ref. 24]. ■

COROLLARY 2 *Let $P_{ijkl} \equiv E(\alpha_i\alpha_j\alpha_k\alpha_l) - E(\alpha_i\alpha_k)E(\alpha_j\alpha_l)$ and define*

$$M_i = \max_{(i-m)h \le x \le ih} f(x)$$

and

$$M_0 = \max_{0 \le x \le 2(m-1)h} f(x).$$

If $k, l \le 2m-2$ and $i, j \ge 3m-2$, then

$$0 \le P_{ijkl} \le 10M_0M_i\frac{h^2}{N} + O\left(\frac{h}{N^2}\right).$$

Proof We note first that the subscript range for i, j, k and l have the property that the basis function $B_i(x)$ and $B_j(x)$ do not overlap the basis functions $B_k(x)$ and $B_l(x)$. So expressions like $E(B_i(x)B_k(x))$, for example, are zero. We then observe that P_{ijkl} is nonnegative by the Cauchy–Schwarz inequality. Finally, since $0 \le B_i(x) \le 1$ and $\int_0^\infty B_i(x)\,dx \le h$, then by

Lemma 7,

$$
\begin{aligned}
P_{ijkl} &= E(\alpha_i\alpha_j\alpha_k\alpha_l) - E(\alpha_i\alpha_k)E(\alpha_j\alpha_l) \\
&\le \frac{1}{N}|-8E(B_i)E(B_j)E(B_k)E(B_l) + E(B_iB_j)E(B_k)E(B_l) \\
&\quad + E(B_i)E(B_j)E(B_kB_l)| + O\left(\frac{h}{N^2}\right) \\
&\le \frac{10E(B_i)E(B_k)}{N} + O\left(\frac{h}{N^2}\right) \\
&\le \frac{10M_iM_0h^2}{N} + O\left(\frac{h}{N^2}\right).
\end{aligned}
$$

■

In preparation for Lemmas 8 and 9, we define $\beta_i^* = \alpha_{L_i}^* - \bar{\alpha}\alpha_i$ for $i = 1, \ldots, m-1$, $\beta_i^* = -\bar{\alpha}\alpha_i$ for $i \ge m$ and

$$\hat{g}_N(x) = \sum_{i=1}^{\infty} \beta_i^* \frac{B_i^m(x)}{b_i}.$$

Then for $\hat{f}_C(x)$ as defined by equation (3) we have for $I = [0, \infty)$, $\hat{f}_C(x) = \hat{f}_N(x) + \hat{g}_N(x)$. Finally, we define the operator $\Delta f = f - E(f)$ and observe that $\Delta\hat{f}_C = \Delta\hat{f}_N + \Delta\hat{g}_N$.

In Lemmas 8 and 9, we write $B^{(n)}$ for the nth derivative of B^m suppressing the explicit dependence of the order of the spline. These two lemmas are used in the computation of the integrated variance. Their proofs are also straightforward and have been omitted.

LEMMA 8

$$E\left(\left(\Delta\hat{f}_N^{(n)}\right)^2\right)\mathrm{d}x = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}(E(\alpha_i\alpha_j) - E(\alpha_i)E(\alpha_j))\int_0^{\infty}\frac{B_i^{(n)}(x)}{b_i}\frac{B_j^{(n)}(x)}{b_j}\,\mathrm{d}x.$$

LEMMA 9

$$E\left(\left(\Delta\hat{g}_N^{(n)}\right)^2\right) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}(E(\beta_i^*\beta_j^*) - E(\beta_i^*)E(\beta_j^*))\int_0^{\infty}\frac{B_i^{(n)}(x)}{b_i}\frac{B_j^{(n)}(x)}{b_j}\,\mathrm{d}x.$$

6. Integrated squared bias

To begin this section, we state a few final facts concerning the basis functions. First of all, since $\mathrm{d}/\mathrm{d}x\, N^m(x) = N^{m-1}(x) - N^{m-1}(x-1)$, one can show [cf. 25] that

$$\left|\frac{\mathrm{d}^n}{\mathrm{d}x^n}B_i^m(x)\right| \le \frac{2^n}{h^n}.$$

So,

$$\left|\int_a^b \frac{\mathrm{d}^n}{\mathrm{d}x^n}B_i^m(x)\frac{\mathrm{d}^n}{\mathrm{d}x^n}B_j^m(x)\,\mathrm{d}x\right| \le \int_{-\infty}^{\infty}\left|\frac{\mathrm{d}^n}{\mathrm{d}x^n}B_i^m(x)\frac{\mathrm{d}^n}{\mathrm{d}x^n}B_j^m(x)\right|\mathrm{d}x \le \frac{4^n}{h^{2n}}\int_{S_{i,j}}1\,\mathrm{d}x \le \frac{m4^n}{h^{2n-1}},$$

where $S_{i,j}$ is the intersection of the supports of $B_i^m(x)$ and $B_j^m(x)$. We now present the proof of the formula for the integrated squared bias of the estimator $\hat{f}_C^{(n)}$ for $m \ge n+2$.

Proof of Theorem 2 We break the integrated squared bias into three integrals and evaluate them separately.

$$\int_0^\infty \left(E\left(\hat{f}_C^{(n)}(x)\right) - f^{(n)}(x)\right)^2 \mathrm{d}x = \int_0^\infty \left(E\left(\hat{f}_N^{*(n)}(x)\right) - f^{(n)}(x)\right)^2 \mathrm{d}x \tag{7}$$

$$- 2\int_0^\infty \left(E\left(\hat{f}_N^{*(n)}(x)\right) - f^{(n)}(x)\right) E\left(\bar{\alpha}\hat{f}_N^{(n)}(x)\right) \mathrm{d}x \tag{8}$$

$$+ \int_0^\infty \left(E\left(\bar{\alpha}\hat{f}_N^{(n)}(x)\right)\right)^2 \mathrm{d}x. \tag{9}$$

Integrals (7) and (8) and (9) are evaluated in order.

Integral (7): Since $p_i^M = O(h^M)$ and $\mathrm{d}^n/\mathrm{d}x^n B_i^m(x) = O(h^{-n})$, it follows that $p_i^{n+2}(x_l)\mathrm{d}^n/\mathrm{d}x^n B_i^m(x) = O(h^2)$. So, by Lemma 1,

$$E\left(f_N^{*(n)}(x)\right) = f^{(n)}(x_l) + f^{(n+1)}(x_l)(x - x_l) + O(h^2).$$

Since

$$f^{(n)}(x) = f^{(n)}(x_l) + f^{(n+1)}(x_l)(x - x_l) + O(h^2),$$

then

$$\int_0^{(m-1)h} \left(E\left(\hat{f}_N^{*(n)}(x)\right) - f^{(n)}(x)\right)^2 \mathrm{d}x = O(h^5).$$

For $x > (m-1)h$, $f_N^*(x)$ is identical to the uncorrected estimator and hence from the analysis in ref. [12]

$$\int_{(m-1)h}^\infty \left(E\left(\hat{f}_N^{*(n)}(x)\right) - f^{(n)}(x)\right)^2 \mathrm{d}x = h^4 C_{m,n} \int_{(m-1)h}^\infty \left(f^{(n+2)}(x)\right)^2 \mathrm{d}x + O(h^5)$$
$$= h^4 C_{m,n} \int_0^\infty \left(f^{(n+2)}(x)\right)^2 \mathrm{d}x + O(h^5),$$

where $C_{mn} = (m/12)^2$ if $m \geq n+3$ and $C_{mn} = 1/720 + (m/12)^2$ if $m = n+2$ [see ref. 12].

Thus, we conclude that Integral (7) satisfies

$$\int_0^\infty \left(E\left(\hat{f}_N^{*(n)}(x)\right) - f^{(n)}(x)\right)^2 \mathrm{d}x = h^4 C_{m,n} \int_0^\infty \left(f^{(n+2)}(x)\right)^2 \mathrm{d}x + O(h^5).$$

Integral (8): Using linearity and Lemma 7, we observe that

$$\int_0^\infty \left(E\left(\hat{f}_N^{*(n)}(x)\right) - f^{(n)}(x)\right) E\left(\bar{\alpha}\hat{f}_N^{(n)}(x)\right) \mathrm{d}x$$
$$= \left(1 - \frac{1}{N}\right) E(\bar{\alpha}) \int_0^\infty \left(E\left(\hat{f}_N^{*(n)}(x)\right) - f^{(n)}(x)\right) E\left(\hat{f}_N^{(n)}(x)\right) \mathrm{d}x$$
$$+ \frac{1}{N}\int_0^\infty \left(E\left(\hat{f}_N^{*(n)}(x)\right) - f^{(n)}(x)\right)$$
$$\times \sum_{i=1}^\infty \left(\sum_{k=1}^{m-1}\sum_{l=k}^{m+k-1} c_{k,l+1-k} E(B_i(X)B_l(X))\right) \frac{B_i^{(n)}(x)}{b_i} \mathrm{d}x.$$

Since $B_i(x)B_l(x) = 0$ when $|i - l| \geq m$ and $B_i^{(n)}(x) = O(h^{-n})$, it follows that

$$\int_0^\infty \left(E\left(\hat{f}_N^{*(n)}(x)\right) - f^{(n)}(x)\right) \times \sum_{i=1}^{\infty}\left(\sum_{k=1}^{m-1}\sum_{l=k}^{m+k-1} c_{k,l+1-k} E(B_i(X)B_l(X))\right)\frac{B_i^{(n)}(x)}{b_i}\,\mathrm{d}x = O(h^{2-n}),$$

therefore

$$\int_0^\infty \left(E\left(\hat{f}_N^{*(n)}(x)\right) - f^{(n)}(x)\right) E\left(\bar{\alpha}\hat{f}_N^{(n)}(x)\right)\mathrm{d}x = \left(1 - \frac{1}{N}\right)E(\bar{\alpha})\int_0^\infty \left(E\left(\hat{f}_N^{*(n)}(x)\right) - f^{(n)}(x)\right)E\left(\hat{f}_N^{(n)}(x)\right)\mathrm{d}x + O\left(\frac{h^2}{Nh^n}\right).$$

We may also observe that

$$\int_0^\infty \left(E\left(\hat{f}_N^{*(n)}(x)\right) - f^{(n)}(x)\right) E\left(\hat{f}_N^{(n)}(x)\right)\mathrm{d}x = \int_{(m-1)h}^\infty \left(E\left(\hat{f}_N^{*(n)}(x)\right) - f^{(n)}(x)\right)E\left(\hat{f}_N^{(n)}(x)\right)\mathrm{d}x + O(h^3).$$

Therefore, apply Lemma 2 to a typical subinterval to obtain

$$\begin{aligned}
&\int_{(m-1)h}^{mh} \left(E\left(\hat{f}_N^{*(n)}(x)\right) - f^{(n)}(x)\right)E\left(\hat{f}_N^{(n)}(x)\right)\mathrm{d}x \\
&= \int_{(m-1)h}^{mh}\left[\sum_{i=m}^{2m-1} p_i^{n+2}\frac{\mathrm{d}^n}{\mathrm{d}x^n}B_i^m(x) - f^{(n+2)}(\eta_x)\frac{(x - x_m)^2}{2}\right] \times \sum_{j=m}^{2m-1} p_j^n \frac{d^n}{dx^n}B_j^m(x)\,\mathrm{d}x \\
&= \sum_{i=m}^{2m-1}\sum_{j=m}^{2m-1}\frac{\Delta^n p_i^{n+2}}{h^n}\frac{\Delta^n p_j^n}{h^n}\int_{(m-1)h}^{mh} B_i^{m-n}(x)B_j^{m-n}(x)\,\mathrm{d}x \\
&\quad - \sum_{j=m}^{2m-1}\frac{\Delta^n p_j^n}{h^n}\int_{(m-1)h}^{mh} f^{(n+2)}(\eta_x)\frac{(x - x_m)^2}{2}B_j^{m-n}(x)\,\mathrm{d}x \\
&= \sum_{i=m}^{2m-1}\sum_{j=m}^{2m-1}\frac{\Delta^n p_i^{n+2}}{h^{n+2}}\frac{\Delta^n p_j^n}{h^n}\int_{(m-1)h}^{mh}\frac{B_i^{m-n}(x)B_j^{m-n}(x)}{h}\,\mathrm{d}x\, h^3 \\
&\quad - \sum_{j=m}^{2m-1}\frac{\Delta^n p_j^n}{h^n} f^{(n+2)}(\eta_j)\int_{(m-1)h}^{mh}\frac{(x - x_m)^2}{2}\frac{B_j^{m-n}(x)}{h^3}\,\mathrm{d}x\, h^3
\end{aligned}$$

by Theorem 8. Now p_i^n/h^n is $f^{(n)}(\xi_j)$ times a constant or the sum of two such terms where the constant is independent of h. So this last expression can be written as

$$h^2 \sum_k f^{(n)}(\xi_k) f^{(n+2)}(\eta_k) d_k h$$

where d_k is independent not only of h but also of the interval. So conclude by Theorem 7 that

$$\begin{aligned}
&\int_0^\infty \left(E(\hat{f}_N^{*(n)}(x) - f^{(n)}(x)\right) E\left(\hat{f}_N^{(n)}(x)\right) \mathrm{d}x \\
&\quad = \lambda h^2 \int_{(m-1)h}^\infty f^{(n)}(x) f^{(n+2)}(x)\, \mathrm{d}x + O(h^3) \\
&\quad = \lambda h^2 \int_0^\infty f^{(n)}(x) f^{(n+2)}(x)\, \mathrm{d}x + O(h^3)
\end{aligned}$$

for some constant λ. Letting $x_m = (m - 1/2)h$, we can see by Lemma 2 that

$$\lambda = \int_{(m-1)h}^{mh} \left(g_{n+2}^{(n)}(x, x_m) - \frac{(x - x_m)^2}{2} \right) \frac{g_n^{(n)}(x, x_m)}{h^3} \,\mathrm{d}x.$$

When $m = n + 2$, $g_{n+2}^{(n)}(x, x_m) = h^2(2m + 1)/24$ and when $m > n + 2$, $g_{n+2}^{(n)}(x, x_m) = h^2 m/12 + (x - x_m)^2/2$. In either case, we determine that $\lambda = m/12$.

By Lemma 6, we then conclude that

$$\begin{aligned}
&\int_0^\infty \left(E\left(\hat{f}_N^{*(n)}(x)\right) - f^{(n)}(x)\right) E\left(\bar{\alpha} \hat{f}_N^{(n)}(x)\right) \mathrm{d}x \\
&\quad = -\left(\frac{m}{12}\right)^2 h^4 f'(0) \int_0^\infty f^{(n)}(x) f^{(n+2)}(x)\, \mathrm{d}x + O(h^5).
\end{aligned}$$

Integral (9): This integral can be similarly evaluated using the ideas for Integrals (7) and (8) and Lemma 3 to yield

$$\int_0^\infty \left(E\left(\bar{\alpha} \hat{f}_N^{(n)}(x)\right)\right)^2 \mathrm{d}x = h^4 f'(0)^2 \left(\frac{m}{12}\right)^2 \int_0^\infty \left(f^{(n)}(x)\right)^2 \mathrm{d}x + O(h^5).$$

The final result is now easily obtained by adding together the values of Integrals (7), (8) and (9). ■

7. Integrated variance

We now present the integrated variance of the estimator $\hat{f}_C^{(n)}$ for $m \geq 2$. Throughout this section, we use $B_i(x)$ instead of $B_i^m(x)$.

Proof of Theorem 3 We break the integrated variance into three integrals and evaluate them separately. So recalling that $\Delta \hat{f}_C = \Delta \hat{f}_N + \Delta \hat{g}_N$, the integrated variance of $\hat{f}_C^{(n)}$ may be written as

$$\int_0^\infty E\left(\left(\Delta \hat{f}_C^{(n)}\right)^2\right) \mathrm{d}x$$

$$= \int_0^\infty E\left(\left(\Delta \hat{f}_N^{(n)}\right)^2\right) \mathrm{d}x \tag{10}$$

$$+ 2\int_0^\infty E\left(\Delta \hat{f}_N^{(n)} \Delta \hat{g}_N^{(n)}\right) \mathrm{d}x \tag{11}$$

$$+ \int_0^\infty E\left(\left(\Delta \hat{g}_N^{(n)}\right)^2\right) \mathrm{d}x. \tag{12}$$

Integrals (10) and (12) are evaluated first, and these results are then used in evaluating Integral (11).

Integral (10): We first note that since $h/m! = b_1 \leq b_i$ for each i that

$$\left|\int_0^\infty B_i^{(n)}(x) \frac{B_j^{(n)}(x)}{b_j} \mathrm{d}x\right| \leq m! \frac{m4^n}{h^{2n}}.$$

So express (10) as

$$\int_0^{(m-1)h} E\left(\left(\Delta \hat{f}_N^{(n)}\right)^2\right) \mathrm{d}x + \int_{(m-1)h}^\infty E\left(\left(\Delta \hat{f}_N^{(n)}\right)^2\right) \mathrm{d}x.$$

Then, using Lemma 8 for the first step followed by an application of Lemma 7 and the integral mean value theorem, we see that

$$\begin{aligned}
&\int_0^{(m-1)h} E\left(\left(\Delta \hat{f}_N^{(n)}\right)^2\right) \mathrm{d}x \\
&= \int_0^{(m-1)h} \sum_{i=1}^{2(m-1)} \sum_{j=1}^{2(m-1)} (E(\alpha_i \alpha_j) - E(\alpha_i)E(\alpha_j)) \frac{B_i^{(n)}(x)}{b_i} \frac{B_j^{(n)}(x)}{b_j} \mathrm{d}x \\
&= \sum_{i=1}^{2(m-1)} \sum_{j=1}^{2(m-1)} \frac{1}{N}(E(B_i B_j) - E(B_i)E(B_j)) \int_0^{(m-1)h} \frac{B_i^{(n)}(x)}{b_i} \frac{B_j^{(n)}(x)}{b_j} \mathrm{d}x \\
&= \frac{1}{N} \sum_{i=1}^{2(m-1)} \sum_{j=1}^{2(m-1)} E\left(\frac{B_i}{b_i} B_j\right) - E\left(\frac{B_i}{b_i}\right) E(B_j) \int_0^{(m-1)h} B_i^{(n)}(x) \frac{B_j^{(n)}(x)}{b_j} \mathrm{d}x \\
&\leq m! \frac{m4^n}{Nh^{2n}} \sum_{i=1}^{2(m-1)} \sum_{j=1}^{2(m-1)} E\left(\frac{B_i}{b_i} B_j\right) + E\left(\frac{B_i}{b_i}\right) E(B_j) \\
&\leq m! \frac{m4^n}{Nh^{2n}} \sum_{i=1}^{2(m-1)} E\left(\frac{B_i}{b_i}\right) + E\left(\frac{B_i}{b_i}\right) \\
&= m! \frac{2m4^n}{Nh^{2n}} \sum_{i=1}^{2(m-1)} f(\xi_i) = O\left(\frac{1}{Nh^{2n}}\right),
\end{aligned}$$

where ξ_i is a point in $[0,(2m-1)h]$. Additionally, we know from Redner [12] that for $m \geq n+1$,

$$\int_{(m-1)h}^{\infty} E\left(\left(\Delta \hat{f}_N^{(n)}\right)^2\right) \mathrm{d}x = \frac{D_{m,n}}{Nh^{1+2n}} + O\left(\frac{1}{Nh^{2n}}\right).$$

Hence it is apparent that Integral (10) satisfies

$$\int_{0}^{\infty} E\left(\left(\Delta \hat{f}_N^{(n)}\right)^2\right) \mathrm{d}x = \frac{D_{m,n}}{Nh^{1+2n}} + O\left(\frac{1}{Nh^{2n}}\right).$$

To finish the proof, we need only show that Integral (11) is $O\left(\sqrt{h}/Nh^{1+2n}\right)$ and Integral (12) is $O\left(1/Nh^{2n}\right)$ and are hence higher order terms. We begin with Integral (12).

Integral (12): We also write Integral (12) as the sum of two integrals:

$$\int_{0}^{4(m-1)h} E\left(\left(\Delta \hat{g}_N^{(n)}\right)^2\right) \mathrm{d}x + \int_{4(m-1)h}^{\infty} E\left(\left(\Delta \hat{g}_N^{(n)}\right)^2\right) \mathrm{d}x. \tag{13}$$

Then, by Lemma 9, we have that the second integral in equation (13) is

$$\begin{aligned}
&\int_{4(m-1)h}^{\infty} E\left(\left(\Delta \hat{g}_N^{(n)}\right)^2\right) \mathrm{d}x \\
&\quad = \int_{4(m-1)h}^{\infty} \sum_{i=4m-3}^{\infty} \sum_{j=4m-3}^{\infty} (E(\beta_i^* \beta_j^*) - E(\beta_i^*)E(\beta_j^*)) \frac{B_i^{(n)}(x)}{b_i} \frac{B_j^{(n)}(x)}{b_j} \mathrm{d}x \\
&\quad = \sum_{i=4m-3}^{\infty} \sum_{j=i-m+1}^{i+m-1} (E(\alpha_i \bar{\alpha} \alpha_j \bar{\alpha}) - E(\alpha_i \bar{\alpha})E(\alpha_j \bar{\alpha})) \int_{4(m-1)h}^{\infty} \frac{B_i^{(n)}(x)}{h} \frac{B_j^{(n)}(x)}{h} \mathrm{d}x
\end{aligned}$$

since $\bar{\alpha}$ is a linear combination of $\alpha_1, \ldots, \alpha_{2m-2}$ and $B_i(x)$ and $B_j(x)$ do not overlap when $|i-j| \geq m$. If we write $\bar{\alpha} = \sum_{k=1}^{2m-2} d_k \alpha_k$, then

$$\begin{aligned}
E(\alpha_i \bar{\alpha} \alpha_j \bar{\alpha}) - E(\alpha_i \bar{\alpha})E(\alpha_j \bar{\alpha}) &= \sum_{k} \sum_{l} d_k d_l \left(E(\alpha_i \alpha_k \alpha_j \alpha_l) - E(\alpha_i \alpha_k)E(\alpha_j \alpha_l)\right) \\
&= \sum_{k=1}^{2m-2} \sum_{l=1}^{2m-2} d_k d_l P_{ijkl},
\end{aligned}$$

where $P_{ijkl} = E(\alpha_i \alpha_k \alpha_j \alpha_l) - E(\alpha_i \alpha_k)E(\alpha_j \alpha_l)$ as in Corollary 2. Hence,

$$\begin{aligned}
&\int_{4(m-1)h}^{\infty} E\left(\left(\Delta \hat{g}_N^{(n)}\right)^2\right) \mathrm{d}x \\
&\quad = \sum_{i=4m-3}^{\infty} \sum_{j=i-m+1}^{i+m-1} \sum_{k} \sum_{l} d_k d_l P_{ijkl} \int_{4(m-1)h}^{\infty} \frac{B_i^{(n)}(x)}{h} \frac{B_j^{(n)}(x)}{h} \mathrm{d}x \\
&\quad \leq \frac{4^n m K_h}{h^{1+2n}} \cdot \sum_{i=3m-2}^{\infty} \sum_{j=3m-2}^{\infty} \sum_{k} \sum_{l} P_{ijkl}
\end{aligned}$$

where $K_h = \max d_k d_l$ and is $O(1)$.

As each of the terms in P_{ijkl} is $O(h/N)$, then by Lemma 7 and Theorem 6

$$\int_{4(m-1)h}^{\infty} E\left(\left(\Delta\hat{g}_N^{(n)}\right)^2\right) \mathrm{d}x \le \frac{10cM_0 4^n m K_h}{h^{2n}N} \sum_{i=4m-3}^{\infty} M_i h + O\left(\frac{1}{h^{2n}N}\right)$$
$$= O\left(\frac{1}{h^{2n}N}\right),$$

where the positive constant c accounts for the finite number of terms with subscripts j, k and l for each index i.

We now evaluate the rest of Integral (10). By another application of Lemma 9 we have

$$\int_0^{4(m-1)h} E((\Delta\hat{g}_N)^2)\,\mathrm{d}x$$
$$= \int_0^{4(m-1)h} \sum_{i=1}^{5m-5} \sum_{j=1}^{5m-5} (E(\beta_i^*\beta_j^*) - E(\beta_i^*)E(\beta_j^*)) \frac{B_i^{(n)}(x)}{b_i} \frac{B_j^{(n)}(x)}{b_j}\,\mathrm{d}x$$
$$\le \sum_{i=1}^{5m-5} \sum_{j=1}^{5m-5} |E(\beta_i^*\beta_j^*) - E(\beta_i^*)E(\beta_j^*)| \int_0^{4(m-1)h} \frac{B_i^{(n)}(x)}{b_i} \frac{B_j^{(n)}(x)}{b_j}\,\mathrm{d}x$$
$$\le \sum_{i=1}^{5m-5} \sum_{j=1}^{5m-5} |E(\beta_i^*\beta_j^*) - E(\beta_i^*)E(\beta_j^*)| \frac{(m!)^2 m 4^n}{h^{2n+1}}$$
$$= O\left(\frac{1}{Nh^{2n}}\right)$$

since $|E(\beta_i^*\beta_j^*) - E(\beta_i^*)E(\beta_j^*)| = O(h/N)$ and there are only a finite number of terms.

After combining our analysis of the first and second integrals in equation (13) we have that Integral (12) is

$$\int_0^{\infty} E\left(\left(\Delta\hat{g}_N^{(n)}\right)^2\right) \mathrm{d}x = O\left(\frac{1}{Nh^{2n}}\right).$$

Integral (11): By two applications of the Cauchy–Schwarz inequality we have for equation (11) that

$$\left(\int_0^{\infty} |E(\Delta\hat{f}_N^{(n)}\Delta\hat{g}_N^{(n)})|\,\mathrm{d}x\right)^2 \le \left(\int_0^{\infty} \sqrt{E((\Delta\hat{f}_N^{(n)})^2)}\sqrt{E((\Delta\hat{g}_N^{(n)})^2)}\,\mathrm{d}x\right)^2$$
$$\le \int_0^{\infty} E((\Delta\hat{f}_N^{(n)})^2)\,\mathrm{d}x \cdot \int_0^{\infty} E((\Delta\hat{g}_N^{(n)})^2)\,\mathrm{d}x.$$

But we know that Integral (10) satisfies

$$\int_0^{\infty} E((\Delta\hat{f}_N^{(n)})^2)\,\mathrm{d}x = \frac{D_{m,n}}{Nh^{1+2n}} + O\left(\frac{1}{Nh^{2n}}\right)$$

and Integral (12) satisfies

$$\int_0^{\infty} E((\Delta\hat{g}_N^{(n)})^2)\,\mathrm{d}x = O\left(\frac{1}{Nh^{2n}}\right),$$

therefore

$$\int_0^\infty E((\Delta \hat{f}_N^{(n)})^2)\,\mathrm{d}x \cdot \int_0^\infty E((\Delta \hat{g}_N^{(n)})^2)\,\mathrm{d}x = O\left(\frac{1}{N^2h^{1+4n}}\right) = O\left(\frac{h}{(Nh^{1+2n})^2}\right).$$

Taking the square root of this term gives us that Integral (11) is

$$\int_0^\infty |E(\Delta \hat{f}_N^{(n)} \Delta \hat{g}_N^{(n)})|\,\mathrm{d}x = O\left(\frac{\sqrt{h}}{Nh^{1+2n}}\right),$$

and this is the desired result.

Finally, by combining the analysis of Integrals (10), (11) and (12) we observe that the integrated variance of the nth derivative of the estimator $\hat{f}_C^{(n)}$ with $m \geq 2$ is

$$\int_0^\infty E((\Delta \hat{f}_C^{(n)}(x))^2)\,\mathrm{d}x = \frac{D_{m,n}}{Nh^{1+2n}} + O\left(\frac{\sqrt{h}}{Nh^{1+2n}}\right) + O\left(\frac{1}{Nh^{2n}}\right).$$

■

8. Discussion

In figure 1, we see the graph of the B-spline nonparametric density estimate without boundary correction (represented as a solid line) as compared to the true density function (represented as a dashed line). One can see the effects on the boundary where the integrated square bias is only $O(h^2)$. The estimate was made using a quadratic ($m = 3$) B-spline density estimate with $h = 0.2$ and with 20,000 data points. We used a large number of points so that the estimate would be very smooth so that one could easily see the bias at the endpoints. In figure 2, we have added the endpoint correction as described in this paper. We see that the affects of the bias are greatly reduced.

Experience with the B-spline density estimator with endpoint correction has shown us, however, that the variation at the endpoints is quite large for smaller sample sizes. Furthermore, the problem is exasperated by increases in the order m of the spline.

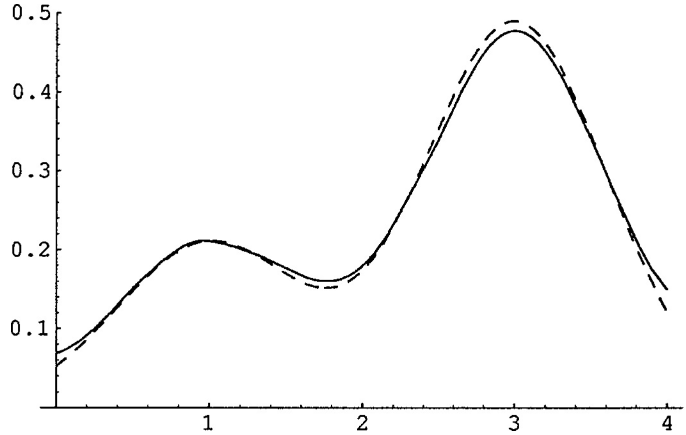

Figure 1. Estimated density without end correction. Dashed line is the true density, solid line is the estimated density ($m = 3$, $h = 0.2$ and $N = 20{,}000$).

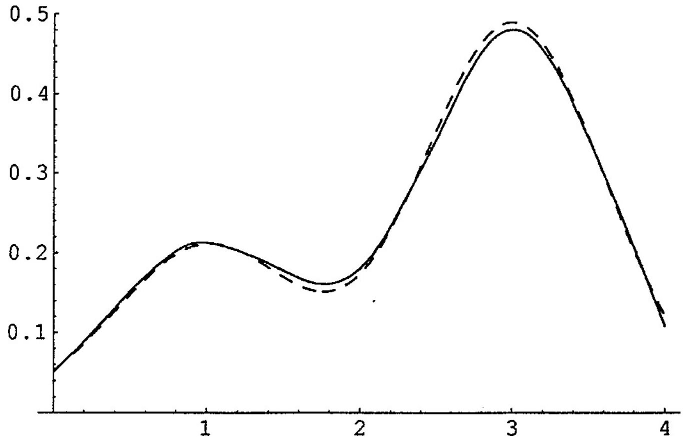

Figure 2. Estimated density with end correction. Dashed line is the true density, solid line is the estimated density ($m = 3$, $h = 0.2$ and $N = 20{,}000$).

We note by looking at Appendix B that some of the correction coefficients for larger values of m are quite large and this too hints that there may still be practical problems related to small sample sizes at the endpoints of the intervals. We also note that while Integral (10) in the proof of Theorem 3 is asymptotically small, the constant in the bound grows large rapidly with m. Finally, one observes that the number of sample points that fall in the support of the basis functions that are truncated at the boundary will be relatively small. In particular, the number of data points that lie in the support of B_1 will be very small.

A number of possible solutions present themselves. The first comes from noting that it may not be necessary to correct all $m - 1$ derivatives in order to get a good estimate of the probability density function. In fact, only the value of the estimate and its first derivative may need to be corrected in order to get a good estimate of the probability density function. This, however, leaves us with too few equations to define the correction coefficients and further study is needed to determine the best way to resolve this ambiguity. Another possible approach is to use equally space knots in the interior of the domain, but to have multiple knots at the endpoint(s). This approach is similar to the beta kernel estimate of Chen [18]. This approach needs to be investigated numerically and a theory for the rate of convergence developed. Finally, these piecewise polynomial B-spline methods need to be compared through simulations with local polynomial density estimate as presented by Cheng *et al.* [20].

The ideas presented in this paper can be used to modify the product tensor B-spline density estimate of [13], but as of this date the theory for these estimators has not yet been investigated. This is an important step as the ultimate goal of this research is to develop nonparametric density estimators suitable for multidimensional problem. Accurate estimates of multidimension probability density functions will require large sample sizes and B-spline density estimators are well suited for handling very large data sets.

References

[1] Redner, R.A. and Gehringer, K.R., 1994, Function estimation using partitions of unity. *Communications in Statistics–Theory and Methods*, **23**(7), 2059–2076.

[2] Sham, P., 1998, *Statistics in Human Genetics* (Arnold).

[3] Atiligan, T. and Bozdogan, H., 1992, Convergence properties of mle's and asymptotic simultaneous confidence intervals in fitting cardinal B-splines for density estimation. *Statistics and Probability Letters*, **13**, 89–98.

[4] Bennett, J.O., 1974, *Estimation of Multivariate Probability Density Functions Using B-Splines* (Houston: PhD thesis, Rice University).

[5] Boneva, L.I., Kendall, D.G. and Stefanov, I., 1971, Spline transformation: Three new diagnostic aids for the statistical data-analyst (with discussion). *Journal of the Royal Statistical Society, Series B, Methodological*, **33**, 1–70.

[6] de Montricher, G.F., Tapia, R.A. and Thompson, J.R., 1975, Nonparametric maximum likelihood estimation of probability densities by penalty function methods. *Annals of Statistics*, **3**, 1329–1348.
[7] Lii, K.-S., 1978, A global measure of a spline density estimate. *Annals of Statistics*, **6**(5), 1138–1148.
[8] Lii, K.-S. and Rosenblatt, M., 1975, Asymptotic behavior of a spline estimate of a density function. *Computers and Mathematics with Applications*, **1**, 223–235.
[9] Stone, C.J., 1994, The use of polynomial splines and their tensor products in multivariate function estimation (with discussion). *Annals of Statistics*, **22**, 117–184.
[10] Wahba, G., 1975, Interpolating spline methods for density estimation. I. Fixed knots. *Annals of Statistics*, **3**(1), 30–48.
[11] Gehringer, K.R. and Redner, R.A., 1992, Nonparametric density estimation using tensor product splines. *Communications in Statistics–Simulation and Computation*, **21**(3), 849–878.
[12] Redner, R.A., 1999, Convergence rates for uniform B-spline density estimators part I: One-dimension. *Siam Journal on Scientific Computing*, **20**(6), 1929–1953.
[13] Redner, R.A., 2000, Convergence rates for uniform B-spline density estimators part II: Multiple dimensions. *Nonparametric Statistics*, **12**, 753–777.
[14] Schuster, E.F., 1985, Incorporating support constraints into nonparametric estimators of densities. *Communications in Statist – Theory and Methods*, **14**, 1123–1136.
[15] Marron, J.S. and Ruppert, D., 1994, Transformations to reduce boundary bias in kernel density estimation. *Journal of the Royal Statistical Society Series B*, **56**(4), 653–671.
[16] Cowling, A. and Hall, P., 1996, On pseudodata methods for removing boundary effects in kernel density estimation. *Journal of the Royal Statistics Society Series B*, **58**(3), 551–563.
[17] Djojosugito, R.A. and Speckman, P.L., 1992, Boundary bias correction in nonparametric density estimation. *Communications in Statistics Part A, Theory Methods*, **21**, 69–88.
[18] Chen, S.X., 1999b, Beta kernel estimator for density functions. *Computational Statistics and Data Analysis*, **31**, 131–145.
[19] Müller, H.G., 1991, Smooth optimal kernel estimators near endpoints. *Biometrika*, **78**, 521–530.
[20] Cheng, M., Fan, J. and Marron, J.S., 1997, On automatic boundary corrections. *The Annals of Statistics*, **25**(4), 1691–1708.
[21] Bartels, R., Beatty, J. and Barsky, B., 1987, *An Introduction to Splines for use in Computer Graphics and Geometric Modeling* (Morgan Kaufmann Publishers).
[22] Schumaker, L.L., 1981, *Spline Functions: Basic Theory* (New York: John Wiley & Sons).
[23] Wolfram, S., 1996, *The Mathematica Book* (Cambridge University Press).
[24] Masri, R.M., 1997, Convergence rates for modified linear B-spline density estimators with semi-infinite and bounded support. Master's thesis, The University of Tulsa.
[25] Chui, C.K., 1992, *An Introduction to Wavelets* (Academic Press, Inc.).
[26] Davis, P.J., 1975, *Interpolation and Approximation* (Dover Publications, Inc.).

Appendix A

The zeros of a real valued function

The following theorem can be easily proved using the ideas concerning the zeros of orthogonal polynomials [see ref. 26, page 236].

THEOREM A1 *Suppose that f is not identically zero on the interval $[a, b]$ and let $f \in C^1[a, b]$ satisfy $\int_a^b f(x)x^j \, \mathrm{d}x = 0$ for $j = 0, 1, \ldots, m-1$. Then f has at least m changes of sign on the interval $[a, b]$.*

COROLLARY A1 *Let $\{B_j^m(x)\}_{j=-\infty}^{\infty}$ be B-spline basis functions based on uniformly spaced knots. If the support of each B_j^m overlaps the interval $[a, b]$ for $j = 1, 2, \ldots, m$, then for $m \geq 1$ the matrix*

$$A = \begin{pmatrix} \int_a^b B_1^m(x)\,\mathrm{d}x & \int_a^b B_2^m(x)\,\mathrm{d}x & \cdots & \int_a^b B_m^m(x)\,\mathrm{d}x \\ \int_a^b x B_1^m(x)\,\mathrm{d}x & \int_a^b x B_2^m(x)\,\mathrm{d}x & \cdots & \int_a^b x B_m^m(x)\,\mathrm{d}x \\ \vdots & \vdots & \vdots & \vdots \\ \int_a^b x^{m-1} B_1^m(x)\,\mathrm{d}x & \int_a^b x^{m-1} B_2^m(x)\,\mathrm{d}x & \cdots & \int_a^b x^{m-1} B_m^m(x)\,\mathrm{d}x \end{pmatrix}$$

is nonsingular.

Proof The result is clearly true when $m = 1$. So consider the case that $m \geq 2$ and assume that A is singular. As the columns are linearly dependent there are constants $c_1, c_2, \ldots, c_m$, not all of which are zero, so that

$$\int_a^b \sum_{j=1}^{m} c_j B_j(x) x^k \, \mathrm{d}x = 0 \quad \text{for } k = 0, 1, \ldots, m-1.$$

Let $f(x) = \sum_{j=1}^{m} c_j B_j(x)$, then when $m \geq 3$, f is continuously differentiable and by Theorem A1, f changes sign at least m times. When $m = 2$, f is piecewise linear and if f changes sign at some point x_1, then clearly $g(x) = f(x)/(x - x_1)$ can be extended to a continuous function on $[a, b]$. So, by the argument made in Theorem A1, we can see that f changes sign at least twice. However, by Theorem 6.2 of Chui [25], f changes sign at most $m - 1$ times. Hence A is nonsingular for all $m \geq 1$. ■

Proof of Theorem 1 The system of equations

$$\sum_{j=i}^{m+i-1} c_{i,j+1-i} \alpha_{j,k} = a_{i,k} b_i - \alpha_{i,k} \quad \text{for } i = 1, \ldots, m-1 \text{ and } k = 0, \ldots, m-1 \tag{A1}$$

can be thought of as $m - 1$ independent subsystems, with one subsystem for each i, $1 \leq i \leq m - 1$. For each i, the subsystem of equations can be written in the following matrix form and has a unique solution by Corollary A1:

$$\begin{pmatrix} \int_0^\infty B_i^m(x)\,\mathrm{d}x & \int_0^\infty B_{i+1}^m(x)\,\mathrm{d}x & \cdots & \int_0^\infty B_{i+m-1}^m(x)\,\mathrm{d}x \\ \vdots & \vdots & \vdots & \vdots \\ \int_0^\infty \frac{x^k}{k!} B_i^m(x)\,\mathrm{d}x & \int_0^\infty \frac{x^k}{k!} B_{i+1}^m(x)\,\mathrm{d}x & \cdots & \int_0^\infty \frac{x^k}{k!} B_{i+m-1}^m(x)\,\mathrm{d}x \\ \vdots & \vdots & \vdots & \vdots \\ \int_0^\infty \frac{x^{m-1}}{(m-1)!} B_i^m(x)\,\mathrm{d}x & \int_0^\infty \frac{x^{m-1}}{(m-1)!} B_{i+1}^m(x)\,\mathrm{d}x & \cdots & \int_0^\infty \frac{x^{m-1}}{(m-1)!} B_{i+m-1}^m(x)\,\mathrm{d}x \end{pmatrix}$$

$$\times \begin{pmatrix} c_{i,1} \\ c_{i,2} \\ \vdots \\ c_{i,m} \end{pmatrix} = \begin{pmatrix} a_{i,0} b_i - \alpha_{i,0} \\ \vdots \\ a_{i,k} b_i - \alpha_{i,k} \\ \vdots \\ a_{i,m-1} b_i - \alpha_{i,m-1} \end{pmatrix}.$$

■

Appendix B

Correction coefficients and variance constants

Table B1. Correction coefficients $c_{i,j}$ in Theorem 1 for $m = 4$.

$c_{i,j}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$
$i = 1$	38.7179	−5.05425	1.12413	−0.163411
$i = 2$	3.30915	−2.84982	1.36141	−0.284906
$i = 3$	0.175755	−0.330564	0.214622	−0.052489

Table B2. Correction coefficients $c_{i,j}$ in Theorem 1 for $m = 5$.

$c_{i,j}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$
$i = 1$	302.154	−21.6381	3.90324	−0.771241	0.0904133
$i = 2$	27.4009	−14.3729	7.04843	−2.39765	0.381747
$i = 3$	1.80477	−3.03309	2.53541	−1.13894	0.212642
$i = 4$	0.0731975	−0.189871	0.196904	−0.0996393	0.0200189

Table B3. Correction coefficients $c_{i,j}$ in Theorem 1 for $m = 6$.

$c_{i,j}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
$i = 1$	2387.97	−99.2459	13.1787	−2.50337	0.433611	−0.0424649
$i = 2$	239.398	−78.355	31.5102	−11.8056	3.10517	−0.395185
$i = 3$	14.8973	−17.683	14.9562	−8.61075	2.92148	−0.436207
$i = 4$	0.919271	−2.36885	2.9319	−2.08417	0.802788	−0.130181
$i = 5$	0.0263641	−0.0886179	0.128433	−0.100192	0.0409746	−0.00692504

Table B4. Integrated variance constants $D_{m,n}$ in Theorems 3 and 5.

Deriv.	$m = 2$	$m = 3$	$m = 4$	$m = 5$	$m = 6$
$n = 0$	0.5	0.396528	0.342609	0.306746	0.280334
$n = 1$		0.402778	0.250972	0.179099	0.136768
$n = 2$			0.569444	0.312018	0.197559
$n = 3$				0.939393	0.475039
$n = 4$					1.66292

Appendix C

Let $\Omega \equiv -\sum_{i=1}^{m-1} \sum_{j=i}^{m+i-1} c_{i,j+1-i}\alpha_{j,1}/h^2$, where $\{\{c_{i,j}\}_{i=1}^{m-1}\}_{j=1}^{m}$ satisfy equation (A1). Through a straightforward but elaborate computation, we will prove that $\Omega = m/12$. For additional verification, the result has also been tested for $2 \le m \le 30$ by direct integration using Mathematica [23].

LEMMA C1

$$\Omega = \sum_{i=0}^{m} \int_{m-i}^{m} \left(y - \frac{m}{2}\right) N^m(y)\,\mathrm{d}y.$$

Proof Observe that $a_{i,1} = \int_{-\infty}^{\infty} x B_i^m(x)/h\,\mathrm{d}x$ is the mean value of a random variable with density $B_i^m(x)/h$. Since this density is symmetric about its mean this equals the mode of the distribution which is $(i - (m/2))h$. Hence

$$\begin{aligned}
\Omega &= \frac{-1}{h^2} \sum_{i=1}^{m-1} \sum_{j=i}^{m+i-1} c_{i,j+1-i}\alpha_{j,1} \\
&= \frac{-1}{h^2} \sum_{i=1}^{m-1} (a_{i,1} b_i - \alpha_{i,1}) \\
&= \frac{-1}{h^2} \sum_{i=1}^{m-1} h\left(i - \frac{m}{2}\right) b_i - \alpha_{i,1} \\
&= \frac{-1}{h^2} \sum_{i=1}^{m-1} \left(h\left(i - \frac{m}{2}\right) \int_0^{\infty} B_i^m(x)\,\mathrm{d}x - \int_0^{\infty} x B_i^m(x)\,\mathrm{d}x \right) \\
&= \frac{-1}{h^2} \sum_{i=1}^{m-1} \int_0^{\infty} \left(h\left(i - \frac{m}{2}\right) - x \right) B_i^m(x)\,\mathrm{d}x \\
&= \frac{-1}{h} \sum_{i=1}^{m-1} \int_0^{\infty} \left(i - \frac{m}{2} - \frac{x}{h}\right) N^m\left(\frac{x}{h} + m - i\right) \mathrm{d}x \\
&= \sum_{i=1}^{m-1} \int_{m-i}^{\infty} \left(y - \frac{m}{2}\right) N^m(y)\,\mathrm{d}y \\
&= \sum_{i=0}^{m} \int_{m-i}^{m} \left(y - \frac{m}{2}\right) N^m(y)\,\mathrm{d}y.
\end{aligned}$$

So define

$$p(t) = \sum_{i=0}^{m} \chi_{[m-i,m]} N^m(t),$$

and observe that

$$\Omega = \int_0^{\infty} \left(t - \frac{m}{2}\right) p(t)\,\mathrm{d}t$$

and that

$$\int_0^{\infty} p(t)\,\mathrm{d}t = \sum_{i=0}^{m} \int_{m-i}^{m} N^m(t)\,\mathrm{d}t = \frac{m+1}{2}$$

by symmetry. So Ω can be determined using the Laplace transform and the fact that

$$\frac{\mathrm{d}}{\mathrm{d}s} \mathcal{L}(p(t))(s)|_{s=0} = -\int_0^{\infty} t p(t)\,\mathrm{d}t.$$

■

The following result is a straightforward calculation and the details have been omitted.

LEMMA C2

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{L}(p(t)) = A(s) + B(s),$$

where

$$A(s) = \sum_{i=0}^{m} \sum_{k=m-i+1}^{m} (-1)^k \binom{m}{k} \frac{e^{-ks}}{s^m}$$

and

$$B(s) = \sum_{i=0}^{m} \sum_{k=0}^{m-i} (-1)^k \binom{m}{k} \sum_{p=0}^{m-1} \frac{e^{-s(m-i)}}{p! s^{m-p}} (m-i-k)^p.$$

LEMMA C3

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{L}(p(t))|_{s=0} = \frac{1}{(m+1)!} \sum_{i=0}^{m} \sum_{k=0}^{m} (-1)^k \binom{m}{k} (-k)^{m+1}$$
$$+ \frac{1}{(m+1)!} \sum_{i=0}^{m} \sum_{k=0}^{m-i} (-1)^k \binom{m}{k} \{-(m-i-k)^{m+1}$$
$$+ (m+1)(m-i)(m-i+k)^m\}.$$

Proof We begin by noting that, since $p(t)$ is piecewise continuous and has bounded support, $\mathcal{L}(p(t))$ has derivatives of all orders on $[0, \infty)$. We also note that for any sufficiently smooth function $g(s)$ that if

$$g(s) = \frac{g^{(m+1)}(0) s^{m+1}}{(m+1)!} + O(s^{m+2})$$

then

$$\frac{\mathrm{d}}{\mathrm{d}s} \frac{g(s)}{s^m}|_{s=0} = \frac{g^{m+1}(0)}{(m+1)!}.$$

So we evaluate

$$\frac{\mathrm{d}}{\mathrm{d}s} s^m A(s)|_{s=0} = \frac{1}{(m+1)!} \frac{d^{m+1}}{ds^{m+1}} \sum_{i=0}^{m} \sum_{k=m-i+1}^{m} (-1)^k \binom{m}{k} e^{-ks}|_{s=0} \tag{A2}$$
$$= \frac{1}{(m+1)!} \sum_{i=0}^{m} \sum_{k=m-i+1}^{m} (-1)^k \binom{m}{k} (-k)^{m+1}.$$

To evaluate the second term we note that

$$\frac{\mathrm{d}^m}{\mathrm{d}s^m} s^p e^{as}|_{s=0} = \binom{m}{p} p! a^{m-p}$$

for $m \geq p$. Hence

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}s} s^m B|_{s=0} \\
&= \frac{1}{(m+1)!} \frac{d^{m+1}}{ds^{m+1}} \sum_{i=0}^{m} \sum_{k=0}^{m-i} (-1)^k \binom{m}{k} \sum_{p=0}^{m-1} \frac{s^p e^{-s(m-i)}}{p!} (m-i-k)^p \\
&= \frac{1}{(m+1)!} \sum_{i=0}^{m} \sum_{k=0}^{m-i} (-1)^k \binom{m}{k} \sum_{p=0}^{m-1} \frac{1}{p!} \binom{m+1}{p} p! \left(-(m-i)^{m+1-p}\right) (m-i-k)^p \\
&= \frac{1}{(m+1)!} \sum_{i=0}^{m} \sum_{k=0}^{m-i} (-1)^k \binom{m}{k} \left\{ \sum_{p=0}^{m+1} \binom{m+1}{p} \left((-(m-i))^{m+1-p} (m-i-k)^p\right) \right. \\
&\quad \left. -(m-i-k)^{m+1} + (m+1)(m-i)(m-i-k)^m \right\} \\
&= \frac{1}{(m+1)!} \sum_{i=0}^{m} \sum_{k=0}^{m-i} (-1)^k \binom{m}{k} \left\{ (-k)^{m+1} - (m-i-k)^{m+1} \right. \\
&\quad \left. + (m+1)(m-i)(m-i-k)^m \right\} \qquad \text{(A3)}
\end{aligned}$$

Combining equations (A2) and (A3) we obtain the final result. ■

LEMMA C4

$$\frac{\mathrm{d}}{\mathrm{d}s} \mathcal{L}(p(t))|_{s=0} = -\frac{m}{3} - \frac{m^2}{4}.$$

Proof This result is obtained by simplifying the expression obtained in Lemma C3. We begin by noting that

$$\begin{aligned}
\frac{1}{(m+1)!} \sum_{i=0}^{m} \sum_{k=0}^{m} (-1)^k \binom{m}{k} (-k)^{m+1} &= -\frac{m+1}{(m+1)!} \sum_{k=0}^{m} (-1)^{m-k} \binom{m}{k} k^{m+1} \\
&= -\frac{m+1}{(m+1)!} \sum_{k=0}^{m} (-1)^{m-k} \binom{m}{k} (j+k)^{m+1}|_{j=0} \\
&= -\frac{m+1}{(m+1)!} \Delta^m j^{m+1}|_{j=0} \\
&= -\frac{m+1}{(m+1)!} (m+1)! \left(j + \frac{m}{2}\right)|_{j=0} \\
&= -(m+1)\left(\frac{m}{2}\right).
\end{aligned}$$

Secondly, we observe that

$$
\begin{aligned}
&\frac{-1}{(m+1)!}\sum_{i=0}^{m}\sum_{k=0}^{m-i}(-1)^k\binom{m}{k}(m-i-k)^{m+1}\\
&\quad=\frac{-1}{(m+1)!}\sum_{k=0}^{m}(-1)^k\binom{m}{k}\sum_{i=0}^{m-k}(m-i-k)^{m+1}\\
&\quad=\frac{-1}{(m+1)!}(-1)^m\sum_{k=0}^{m}(-1)^{m-k}\binom{m}{k}\sum_{i=0}^{m-(j+k)}(m-i-(j+k))^{m+1}|_{j=0}\\
&\quad=\frac{-1}{(m+1)!}(-1)^m\Delta^m\sum_{i=0}^{m-j}(m-i-j)^{m+1}|_{j=0}\\
&\quad=\frac{-1}{(m+1)!}(-1)^m\Delta^{m-1}\left[\sum_{i=0}^{m-j-1}(m-i-(j+1))^{m+1}-\sum_{i=0}^{m-j}(m-i-j)^{m+1}\right]\Bigg|_{j=0}\\
&\quad=\frac{-1}{(m+1)!}(-1)^m\Delta^{m-1}(-1)(m-j)^{m+1}|_{j=0}\\
&\quad=\frac{-1}{(m+1)!}\Delta^{m-1}(j-m)^{m+1}|_{j=0}\\
&\quad=\frac{-1}{(m+1)!}\Delta^{m-1}\left(j+\frac{-m-1}{2}-\frac{m-1}{2}\right)^{m+1}\Bigg|_{j=0}\\
&\quad=\frac{-1}{(m+1)!}(m+1)!\left(\frac{m-1}{24}+\frac{(j+(-m-1)/2)^2}{2}\right)\Bigg|_{j=0}\\
&\quad=-\frac{m-1}{24}-\frac{(m+1)^2}{8}.
\end{aligned}
$$

Finally, we consider the expression

$$
\begin{aligned}
&\frac{1}{(m+1)!}\sum_{k=0}^{m}(-1)^k\binom{m}{k}\sum_{i=0}^{m-k}(m+1)(m-i)(m-i-k)^m\\
&\quad=\frac{(-1)^m}{m!}\sum_{k=0}^{m}(-1)^{m-k}\binom{m}{k}\sum_{i=0}^{m-(j+k)}(m-i)(m-i-(j+k))^m|_{j=0}\\
&\quad=\frac{(-1)^m}{m!}\Delta^m\sum_{i=0}^{m-j}(m-i)(m-i-j)^m|_{j=0}.
\end{aligned}
$$

Observe that

$$
\begin{aligned}
&\Delta \sum_{i=0}^{m-j} (m-i)(m-i-j)^m \\
&\quad = \sum_{i=0}^{m-(j+1)} (m-i)(m-i-(j+1))^m - \sum_{i=0}^{m-j} (m-i)(m-i-j)^m \\
&\quad = -m(m-j)^m + \sum_{i=0}^{m-j-1} (m-i-j-1)^m.
\end{aligned}
$$

Therefore, we consider two terms,

$$\frac{(-1)^m}{m!} \Delta^{m-1} (-m(m-j))^m |_{j=0}$$

and

$$\frac{(-1)^m}{m!} \Delta^{m-1} \sum_{i=0}^{m-j-1} (m-i-j-1)^m |_{j=0}.$$

The first of these terms is easy to evaluate since

$$
\begin{aligned}
\frac{(-1)^m}{m!} \Delta^{m-1} (-m(m-j)^m)|_{j=0} &= \frac{(-1)^{2m+1}}{m!} m \Delta^{m-1} (j-m)^m)|_{j=0} \\
&= \frac{-1}{m!} m m! \left(j - \frac{m+1}{2} \right) \bigg|_{j=0} \\
&= m \frac{m+1}{2}.
\end{aligned}
$$

To evaluate the second term, we first note that

$$
\begin{aligned}
&\Delta \sum_{i=0}^{m-j-1} (m-i-j-1)^m \\
&\quad = \sum_{i=0}^{m-(j+1)-1} (m-i-(j+1)-1)^m - \sum_{i=0}^{m-j-1} (m-i-j-1)^m \\
&\quad = -m(m-j-1)^m.
\end{aligned}
$$

So

$$\begin{aligned}
&\frac{(-1)^m}{m!}\Delta^{m-1}\sum_{i=0}^{m-j-1}(m-i-j-1)^m|_{j=0}\\
&=\frac{(-1)^{m+1}}{m!}\Delta^{m-2}(m-j-1)^m|_{j=0}\\
&=\frac{-1}{m!}\Delta^{m-2}\left(j-\frac{m-2}{2}-\frac{m}{2}\right)^m\Bigg|_{j=0}\\
&=-\frac{m-2}{24}\left(\frac{(j-(m/2))^2}{2}\right)^2\Bigg|_{j=0}\\
&=-\frac{m-2}{24}-\frac{m^2}{8}.
\end{aligned}$$

Then the argument is complete by observing that

$$-m\frac{m+1}{2}-\frac{m-1}{24}-\frac{(m+1)^2}{8}+m\frac{m+1}{2}-\frac{m-2}{24}-\frac{m^2}{8}=-\frac{m}{3}-\frac{m^2}{4}.$$

■

THEOREM C1 $\Omega=\dfrac{m}{12}$.

Proof This result is now easily obtained since

$$\Omega=-\left(\frac{-m}{3}+\frac{-m^2}{4}\right)-\frac{m}{2}\frac{m+1}{2}=\frac{m}{12}.$$

■