No. 30

Higher order kernel density estimation on the circle

Yasuhito Tsuruta

Masahiko Sagae

December, 2016

Higher order kernel density estimation on the circle

Yasuhito Tsuruta [*] Masahiko Sagae [†]

2016/12/1

Abstract

We defined new kernel functions of pth-order with new moments called for effective moments. the kernel functions can give improvement to reduce the bias. We showed that our kernel functions has the convergent rate of the mean integrated square error $O(n^{-2p/(2p+1)})$ and the asymptotic normality. Moreover, we show two methods to construct higher order kernel density estimators. Our simulation syas that the higher order kernel density estimators is better than the second order kernel density estimator as n is large.

1 Introduction

Given a random sample on the circle $\Theta_1, \dots, \Theta_n \in [-\pi, \pi)$ from an unknown density $f(\theta)$. The kernel density estimator (KDE) on the circle is defined as,

$$\hat{f}_\kappa(\theta) = \frac{1}{n} \sum_{i=1}^{n} K_\kappa(\theta - \Theta_i),$$

where $K_\kappa(\theta)$ is a symmetric kernel function and κ is a concentration parameter, which plays a role of a smoothing parameter and corresponds to the inverse of bandwidth on the real line.

Di Marzio et al.(2011) defined sin-order kernel functions with sin-order moments $\eta_j(K_\kappa) := \int_{-\pi}^{\pi} \sin^j(\theta) K_\kappa(\theta) d\theta$ and derived the property of the mean integrated square error (MISE) for $\hat{f}_\kappa$. They showed that the convergent rate of the MISE for the von Mises (VM) kernel (the second sin-order kernel) is $O(n^{-5/4})$. They described that sin-order kernel functions do

[*]Graguate School of Human and Socio-Environment Studies, Kanazawa University
[†]School of Economics, Kanazawa University

not necessarily yield smaller bias. In other words, the order p of the pth sin-order kernel function does not generally correspond to the convergent rate of the MISE. For example,Tsuruta and Sagae (2016) indicated that the optimal MISE's rate of the wrapped Cauchy kernel (the second sin-order kernel) is $O(n^{-2/3})$. They succeeded in construction of higher-order kernels and improvement of the bias by "Twicing" of the bias reduction technique.

We introduce the new class of pth-order kernel functions with new moments in the next section. The convergent rate of the MISE of $\hat{f}_\kappa$ achieves $O(n^{-2p/(2p+1)})$ with the pth-order kernel K_κ. In addition, the $\hat{f}_\kappa$ also has the asymptotic normality.

Higher-order kernel density estimators are constructed by applying either the additive method by Jones ad Foster (2013) or the multiplicative method by terrell and Scott (1980) to our kernel functions. Our simulation shows that these higher-order KDEs have better properties than the second-order KDE as n is moderate or large.

2 Properties of higher-order kernel density estimators

We redefine the kernel functions for deriving higher-order properties of the KDE $\hat{f}_\kappa$. The kernel function is $K_\kappa(\theta) := C^{-1}(L)L(\kappa\{1-\cos(\theta)\})$ defined by [4], where $C_\kappa(L)$ is the normalizing constant.

Definiton 1. *(Kernel function)*

$K_\kappa(\theta) : [-\pi, \pi) \to \mathbb{R}$ *satisfies with the following five conditions,*

(i) It is defined as $K_\kappa(\theta) := C_\kappa^{-1}(L)L_\kappa(\theta)$, *where* $L_\kappa(\theta) := L(\kappa\{1-\cos(\theta)\})$ *and* $C_\kappa(L) := \int_{-\pi}^{\pi} L_\kappa(\theta)d\theta$,

(ii) The term $L(r) \to 0$ *as* $r \to \infty$, *where* $r = \kappa\{1-\cos(\theta)\}$,

(iii) Let be $\delta_{\kappa^{1/2},2t}(L) := \int_{-\kappa^{1/2}\pi}^{\kappa^{1/2}\pi} L^2(z^2/2)z^{2t}dz$, $\delta_{2t}(L) := \int_{-\infty}^{\infty} L^2(z^2/2)z^{2t}dz$, *and* $\delta_{2t}(L)$ *has bounded for t=0,1,*

(iv) Let $l \geq 0$ *and* $v \geq 2$ *be even. Define the* l*th moment as,*

$$\mu_l(L) := \int_0^\infty L(r)r^{(l-1)/2}dr,$$

(v) Let be $\mu_{\kappa,l}(L) := \int_0^\kappa L(r)r^{(l-1)/2}dr$. *The* l*th moment* $\mu_l(L)$ *is bounded and* $\mu_{\kappa,l}(L) = \mu_l(L) + O(\kappa^{-(v+2)/2})$ *for any* $0 \leq l \leq v$.

We define the pth-order kernel function.

Definiton 2. *(pth-order kenel function)*

Let $p \geq 2$ be even. We will say that $K_\kappa(\theta)$ is pth-order kernel for $v \geq p+2$, if,

$$\mu_0(L) \neq 0, \quad \mu_l(L) = 0, \quad l = 2, 4, \cdots, p-2 \quad \text{and} \quad \mu_l(L) \neq 0 \quad l = p.$$

We prove higher-order properties of $\hat{f}_\kappa$ by using the later lemmas 1-3.

Lemma 1. *Let be $g_j(r/\kappa) := \{2 - r/\kappa\}^{(j-1)/2}$ for $j \geq 0$. Then, $C_\kappa(L)$ is given as,*

$$C_\kappa(L) = 2 \sum_{m=0}^{v/2} \frac{g_0^{(m)}(0)}{m!} \kappa^{-(2m+1)/2} \mu_{2m}(L) + O(\kappa^{-(v+1)/2}). \tag{1}$$

If K_κ is pth-order kernel, then the term $C_\kappa(L)$ is reduced to,

$$C_\kappa(L) = \kappa^{-1/2} 2^{1/2} \mu_0(L) + O(\kappa^{-(p+1)/2}). \tag{2}$$

See Appendix-A for the details.

Lemma 2. *Put $\alpha_j(K_\kappa) := \int_{-\pi}^{\pi} K_\kappa(\theta)\theta^j d\theta$ and $a_s := (2s-2)!!/\{(2s-1)!!s\}$. Let $z \geq 0$ be even and $2t \leq z \leq v$. Then, the term $\alpha_{2t}(K_\kappa)$ is given as,*

$$\alpha_{2t}(K_\kappa) = 2C_\kappa^{-1}(L)\kappa^{-1/2} \sum_{q=t}^{z/2} \sum_{m=0}^{z/2-q} \kappa^{-(q+m)} A_q(z,t)(m!)^{-1} g_{2q}^{(m)}(0) \mu_{2(q+m)}(L) + O(\kappa^{-(z+2)/2}), \tag{3}$$

where,

$$A_q(z,t) := \sum_{\sum_{s=1}^{z/2} t_s = t, \quad \sum_s^{z/2} s t_s = q} \frac{t!}{t_1! t_2! \cdots t_{z/2}!} \prod_{l=1}^{z/2} a_l^{t_l}.$$

If K_κ is pth-order kernel, then the equation (3) is reduced to,

$$\alpha_{2t}(K_\kappa) = b_{p,2t} \mu_0^{-1}(L) \mu_p(L) \kappa^{-p/2} + O(\kappa^{-(p+2)/2}) \quad 0 < j \leq p,$$

where,

$$b_{p,2t} = 2^{1/2} \sum_{q=t}^{p/2} A_q(p,t)(\{p/2-q\}!)^{-1} g_{2q}^{(p/2-q)}(0).$$

The term $\alpha_{p+2}(K_\kappa) = O(\kappa^{-(p+2)/2})$ is shown by (3).

See Appendix-B for the details.

Lemma 3. *Let be* $R(g) := \int_{-\pi}^{\pi} g^2(\theta)d\theta$. *The term* $\delta_{\kappa^{1/2},2t}(L) = \delta_{2t}(L) + o(1)$ *by (iii) leads to,*

$$R(K(u)u^t) := \kappa^{-(2t-1)/2}[d_{2t}(L) + o(1)], \tag{4}$$

where $d_{2t}(L) := 2^{-1}\mu_0^{-2}(L)\delta_{2t}(L)$ *and* $d(L) := d_0(L)$.

See Appendix-C for the details.

Put $\mathrm{MSE}[\hat{f}_\kappa(\theta)] := \mathrm{E}[\{\hat{f}_\kappa(\theta) - f(\theta)\}^2]$ and $\mathrm{MISE}[\hat{f}_\kappa] := \int_{-\pi}^{\pi} \mathrm{MSE}[\hat{f}_\kappa(\theta)]d\theta$. The higher-order properties of $\hat{f}_\kappa$ are obtained by lemmas 1-3.

Theorem 1. *(the MISE) Under the conditions:*

(i) Let be $\kappa = \kappa(n)$. *Then,* $\lim_{n\to\infty} \kappa(n) = \infty$,

(ii) $\lim_{n\to\infty} n^{-1}\kappa^{1/2}(n) = 0$,

(iii) f *is (p+2)th differentiable and* $f^{(s)}$, $s = 1, 2, \cdots, p$ *is square-integrable,*

(iv) K_κ *is pth-order kernel functions,*

The MISE is derived as,

$$MISE[\hat{f}_\kappa] = \frac{\mu_p^2(L)}{\mu_0^2(L)} R\left(\sum_{t=1}^{p/2} \frac{b_{p,2t} f^{(2t)}}{2t!}\right)\kappa^{-p} + n^{-1}\kappa^{1/2}d(L) + o(\kappa^{-p} + n^{-1}\kappa^{1/2}). \tag{5}$$

The main terms of (5) *is referred as* $AMISE[\hat{f}_\kappa]$. *The minimizer* κ^* *of* $AMISE[\hat{f}_\kappa]$ *is equal to,*

$$\kappa^* = \left[\frac{2p\mu_p^2(L)R(\sum_{t=1}^{p/2}[b_{p,2t}f^{(2t)}/(2t!)])n}{\mu_0^2(L)d(L)}\right]^{2p/(2p+1)}. \tag{6}$$

Thus, the optimal MISE for KDE $\hat{f}_\kappa$ *with pth order kernel is* $O(n^{-2/(2p+1)})$.

See Appendix-D for the proof.

It is shown that $\hat{f}_\kappa$ has asymptotic normality.

Theorem 2. *(Asymptotic Normality) Let be $\kappa = cn^{\alpha}$. If $\alpha > (2p+1)/2$ and $n \to \infty$,*

$$\sqrt{n\kappa^{-1/2}}[\hat{f}_{\kappa}(\theta) - f(\theta)] \xrightarrow{d} N(0, f(\theta)d(L)).$$

See Appendix-E for the proof.

We explore about MISE properties of the VM kernel function, which is typical kernel on the circle. The main term of the VM kernel function is defined to $L_{\kappa,\mathrm{VM}}(\theta) := \exp[-\kappa\{1 - \cos(\theta)\}]$. The moment of VM kernel is given as $\mu_l(L_{\mathrm{VM}}) = \Gamma(l + 1/2)$ where $\Gamma(1/2) = \sqrt{\pi}$ and $\Gamma(l + 1/2) = (2l-1)!!\sqrt{\pi}/2^l$ for $l \geq 1$. The VM kernel is the second-order kernel since $\mu_2(L_{\mathrm{VM}}) \neq 0$. The VMKDE is represented as $\tilde{f}_{\kappa}^{\mathrm{VM}}(\theta)$. The AMISE of $\tilde{f}_{\kappa}^{\mathrm{VM}}(\theta)$ is given as,

$$\mathrm{AMISE}[\tilde{f}_{\kappa}^{\mathrm{VM}}] = \frac{1}{4}R(f^{(2)})\kappa^{-2} + \frac{\kappa^{1/2}}{2\sqrt{\pi}n}. \tag{7}$$

The AMISE of (7) and the optimal concentration paramator $\kappa_{\mathrm{VM}} = [2\sqrt{\pi}R(f^{(2)})n]^{2/5}$ are equivalent to those of the VM kernel derived by Di Marzio et al. (2011).

We derive two methods for constructing higher-order kernel density estimator. Method-1 (Additive method) is constructed by using the kernel and its derivative. Method-2 (Multiplicative method) is done by combining two KDEs with different concentration parameters. Introduce the notation of the pth-order kernel for the latter discussion. Let pth-order kernel function be $K_{\kappa,[p]} := C_{\kappa}^{-1}(L_{[p]})L_{\kappa,[p]}(\theta)$, where $L_{\kappa,[p]}(\theta) := L_{[p]}(\kappa\{1 - \cos(\theta)\})$.

Method 1. *(Additive method)*

Notice $r = \kappa\{1 - \cos(\theta)\}$. Let be $dL_{[p]}(r)/dr := L'_{[p]}(r)$ and $L_{[p]}(r)$ be differentiable. Then the main term $L^{JF}_{[p+2]}(r)$ of Jones and Foster's (JF) kernel $K^{JF}_{\kappa,[p+2]}$ is defined as,

$$L^{JF}_{[p+2]}(r) := \frac{p+1}{p}L_{[p]}(r) + \frac{2}{p}rL'_{[p]}(r). \tag{8}$$

Then, the JF kernel $K^{JF}_{\kappa,[p+2]}$ is $(p+2)$th-order kernel.

See Appendix-G for the details.

Method 2. *(Multiplicative method)*

The ratio of two second-order KDEs $\hat{f}_\kappa(\theta)$ and $\hat{f}_{\kappa/4}(\theta)$ with different bandwidths constructs the fourth-order Terrell and Scott's (TS) KDE, that is,

$$\hat{f}_\kappa^{TS}(\theta) := \hat{f}_\kappa^{4/3}(\theta)\hat{f}_{\kappa/4}^{-1/3}(\theta).$$

Then,

$$AMISE[\hat{f}_\kappa^{TS}] = R(G)\kappa^{-4} + n^{-1}\kappa^{1/2}D(L), \tag{9}$$

where $G(\theta)$ and $D(L)$ is given by (G.7) *and* (G.12)*, respectively. The minimizer κ_{TS} of* (9) *is given as,*

$$\kappa_{TS}^* := \left[\frac{8R(G)n}{D(L)}\right]^{2/9}. \tag{10}$$

This derives the optimal MISE $= O(n^{-8/9})$.

See Appendix-F for the details.

We generate $\tilde{L}_{\kappa,[4]}^{\mathrm{JF}}(\theta)$ by applying (8) to $L_{\kappa,\mathrm{VM}}(\theta)$ of the VM kernel. Then $\tilde{L}_{\kappa,[4]}^{\mathrm{JF}}(\theta)$ is given as,

$$\tilde{L}_{\kappa,[4]}^{\mathrm{JF}}(\theta) = \{3/2 - \kappa\{1-\cos(\theta)\}\}\exp[-\kappa\{1-\cos(\theta)\}].$$

Let be $\tilde{f}_\kappa^{\mathrm{JF}}(\theta) := n^{-1}\sum_i \tilde{K}_{\kappa,[4]}^{\mathrm{JF}}(\theta)$, where $\tilde{K}_{\kappa,[4]}^{\mathrm{JF}}(\theta) := C_\kappa^{-1}(\tilde{L}_{[4]}^{\mathrm{JF}})\tilde{L}_{\kappa,[4]}^{\mathrm{JF}}(\theta)$. The optimal concentration parameter κ_{JF} is represented as,

$$\kappa_{\mathrm{JF}} = \left[\frac{16\sqrt{\pi}}{3}R\left(\frac{5f^{(2)}+2f^{(4)}}{12}\right)n\right]^{2/9}. \tag{11}$$

This also derives the optimal MISE $= O(n^{-8/9})$.

The optimal concentration parameter κ_{TS} of that $\tilde{f}_\kappa^{\mathrm{TS}}(\theta)$ is derived from two the VMKDEs with different concentration parameters by Method-2. By (10), the parameter κ_{TS} is given by,

$$\kappa_{\mathrm{TS}} = \left[\frac{288}{33-16\sqrt{2}/\sqrt{5}}R\left(\frac{\{f^{(2)}\}^2}{2f} - \frac{5f^{(2)}+2f^{(4)}}{4}\right)n\right]^{2/9}$$

3 Simulation

For practical exemple the optimal concentration parameter depends on a true density f. We use the plug-in rule to estimate it. The plug-in rule assumes that f is the VM density $f_{\mathrm{VM}}(\theta;\tau) := (2\pi)^{-1} I_0(\tau)\exp\{\tau\cos(\theta)\}$, where $I_p(\tau)$ denotes the pth modified Bessel function of the first kind and order 0. The plug-in rule uses the maximum likelihood estimator $\hat{\tau}$ as an initial value. We denote the plug-in rule estimators of κ_{VM}, κ_{JF} and κ_{TS} as $\hat{\kappa}_{\mathrm{VM}}$, $\hat{\kappa}_{\mathrm{JF}}$ and $\hat{\kappa}_{\mathrm{TS}}$, respectively,

$$\begin{aligned}\hat{\kappa}_{\mathrm{VM}} &= [2\sqrt{\pi}\hat{R}_{\hat{\tau}}(f_{\mathrm{VM}}^{(2)})n]^{2/5},\\ \hat{\kappa}_{\mathrm{JF}} &= \left[\frac{16\sqrt{\pi}}{3}\hat{R}_{\hat{\tau}}\left(\frac{5f_{\mathrm{VM}}^{(2)}+2f_{\mathrm{VM}}^{(4)}}{12}\right)n\right]^{2/9},\end{aligned} \tag{12}$$

and,

$$\hat{\kappa}_{\mathrm{TS}} = \left[\frac{288}{33-16\sqrt{2}/\sqrt{5}}\hat{R}_{\hat{\tau}}\left(\frac{\{f_{\mathrm{VM}}^{(2)}\}^2}{2f_{\mathrm{VM}}} - \frac{5f_{\mathrm{VM}}^{(2)}+2f_{\mathrm{VM}}^{(4)}}{4}\right)n\right]^{2/9}.$$

See Appendix-H for the above details.

Our simulation follows the (i)-(iii) procedures :

(i) The VM kernel density estimator $\tilde{f}_\kappa^{\mathrm{VM}}(\theta)$:

- (a) Generate the random sample of the size n distributed as $f_{\mathrm{VM}}(\theta;\tau)$,
- (b) Estimate κ_{VM} by the plug-in rule,
- (c) Calculate $\overline{\mathrm{ISE}}(\tilde{f}_\kappa^{\mathrm{VM}})$, where $\overline{\mathrm{ISE}}(\tilde{f}_\kappa^{\mathrm{VM}})$ is the numerical integration for $\mathrm{ISE} = \int_0^{2\pi}\{\hat{f}(\theta;\rho) - f(\theta)\}^2 d\theta$.
- (d) Repeat (a)-(c) 1000 times and compute $\overline{\mathrm{MISE}}(\tilde{f}_\kappa^{\mathrm{VM}}) = \sum_{i=1}^{1000}\overline{\mathrm{ISE}}_i(\tilde{f}_\kappa^{\mathrm{VM}})/1000$.

(ii) The Jones and Foster kernel density estimator $\tilde{f}_\kappa^{\mathrm{JF}}(\theta)$:
With the same procedure as (a)-(d) of $\tilde{f}_\kappa^{\mathrm{VM}}(\theta)$, compute $\overline{\mathrm{MISE}}(\tilde{f}_\kappa^{\mathrm{JF}})$

(iii) The Terrell and Scott kernel density estimator $\tilde{f}_\kappa^{\mathrm{TS}}(\theta)$:
With the same procedure as (a)-(d) of $\tilde{f}_\kappa^{\mathrm{VM}}(\theta)$, compute $\overline{\mathrm{MISE}}(\tilde{f}_\kappa^{\mathrm{TS}})$.

Table 1 shows that the JFKDE $\tilde{f}_\kappa^{\mathrm{JF}}$ and the TSKDE $\tilde{f}_\kappa^{\mathrm{TS}}$ trend to be superior to the VMKDE $\tilde{f}_\kappa^{\mathrm{VM}}$ under that n is moderate or large . The result is caused by the fact that both $\tilde{f}_\kappa^{\mathrm{JF}}$ and $\tilde{f}_\kappa^{\mathrm{TS}}$ are forth-order, while $\tilde{f}_\kappa^{\mathrm{VM}}$ is second-order. $\tilde{f}_\kappa^{\mathrm{JF}}$ is the best for $\tau \geq 2$, and $\tilde{f}_\kappa^{\mathrm{TS}}$ is the best for $n \geq 100$ and $\tau \leq 1$, however, $\tilde{f}_\kappa^{\mathrm{VM}}$ is the best only if $n = 50$ and $\tau \leq 1$.

Table 1: MISE's are based on 1000 simulated samples of size n =50, 100, 300, 500 and 1000. τ is the concentration parameter of $\mathrm{f_{VM}}$. JFKDE is the Jones and Foster's KDE $\tilde{f}_\kappa^{\mathrm{JF}}$. TSKDE is the Terrell and Scott's $\tilde{f}_\kappa^{\mathrm{TS}}$. VMKDE is the VM kernel density estimator $\tilde{f}_\kappa^{\mathrm{VM}}$.

		$n = 50$	$n = 100$	$n = 200$	$n = 300$	$n = 500$	$n = 1000$
$\tau = 0.5$	JFKDE	0.006899	0.004334	0.002284	0.001567	0.00107	0.000611
	TSKDE	0.006667	**0.00415**	**0.002179**	**0.001494**	**0.000998**	**0.000531**
	VMKDE	**0.006257**	0.004399	0.002413	0.001718	0.001212	0.000699
$\tau = 1$	JFKDE	0.010001	0.005652	0.003199	0.002156	0.001515	0.00085
	TSKEDE	0.009331	**0.005203**	**0.002884**	**0.001937**	**0.001349**	**0.000768**
	VMKDE	**0.009248**	0.005524	0.003319	0.002393	0.001676	0.000997
$\tau = 2$	JFKDE	**0.012996**	**0.00692**	**0.003905**	**0.002731**	**0.001809**	**0.000973**
	TSKEDE	0.013448	0.007196	0.004069	0.002838	0.001875	0.001013
	VMKDE	0.013104	0.0077	0.004719	0.00347	0.002358	0.001394
$\tau = 5$	JFKDE	**0.021106**	**0.011399**	**0.006392**	**0.00442**	**0.00288**	**0.00159**
	TSKEDE	0.022821	0.012223	0.006842	0.004723	0.003078	0.001701
	VMKDE	0.021729	0.012719	0.007763	0.005592	0.003908	0.002275

Acknowledgements

This work was supported by JSPS KAKENHI Grant Numbers JP16K00043,JP24500339, JP16H02790.

References

[1] Di Marzio, M., Panzera, A. and Taylor, C. C., 2009. Local polynomial regression for circular predictors. Statistics & Probability Letters 79, 2066-2075.

[2] Di Marzio, M., Panzera, A. and Taylor, C. C. , 2011. Kernel density estimation on the torus. Journal of Statistical Planning and Inference 141, 2156-2173.

[3] Feller, W. , 1968. An Introduction to Probability Theory and Its Applications. I (Third ed.). John Wiley & Sons, Inc, USA. pp.244.

[4] Hall, P., Watson, G. S. and Cabrera, J. , 1987. Kernel density estimation with spherical data. Biometrika 74, 751-762.

[5] Jones, M. C. and Foster, P.J. , 1993. Generalized jacknifing and higher-order kernels. Journal of Nonparametric Statistics 3, 81-94.

[6] Terrell, G. R. and Scott, D. W. , 1980. On improving convergence rates for nonnegative kernel Density Estimators. The Annals of Statistics 8, 1160-1163.

[7] Tsuruta, Y., and Sagae, M. , 2016. Asymptotic Property of Wrapped Cauchy Kernel Density Estimation on the Circle. Discussion Paper, No.28 in school of Economics, Kanazawa University.

Appendix-A

Proof of Lemma 1. Notice that $g_j(r/\kappa) = \sum_{m=0}^{z/2}(m!)^{-1}g_j^{(m)}(0)(r/\kappa)^m + O(\kappa^{-(z+2)/2})$ and $d\theta/dr = \{r\kappa(2-r/\kappa)\}^{-1/2}$. This leads to,

$$
\begin{aligned}
C_\kappa(L) &= \int_{-\pi}^{\pi} L(\{\kappa(1-\cos(\theta)\})d\theta \\
&= 2\int_0^{2\kappa} L(r)(r\kappa)^{-1/2}\{2-r/\kappa\}^{-1/2} \\
&= 2\sum_{m=0}^{v/2}\frac{g_0^{(m)}(0)}{m!}\kappa^{-(2m+1)/2}\int_0^{2\kappa} L(r)r^{(2m-1)/2}dr + O(\kappa^{-(v+3)/2}) \\
&= 2\sum_{m=0}^{v/2}\frac{g_0^{(m)}(0)}{m!}\kappa^{-(2m+1)/2}\mu_{2\kappa,2m}(L) + O(\kappa^{-(v+3)/2}) \\
&= 2\sum_{m=0}^{v/2}\frac{g_0^{(m)}(0)}{m!}\kappa^{-(2m+1)/2}\mu_{2m}(L) + O(\kappa^{-(v+3)/2}).
\end{aligned}
$$

If K_κ is pth-order kernel, then $C_\kappa(L)$ is given as,

$$
C_\kappa(L) = 2^{1/2}\mu_0(L)\kappa^{-1/2} + O(\kappa^{-(p+1)/2}).
$$

□

Appendix-B

Proof of Lemma 2. Notice that $\sin^2(\theta) = r/\kappa(2-r/\kappa)$ via $r = \kappa\{1-\cos(\theta)\}$. The Taylor expansion of $\theta^2 = \arcsin^2(\{r/\kappa(2-r/\kappa\}^{1/2})$ for $0 \le \theta < \pi/2$ is given as,

$$\theta^2 = \sum_{s=1}^{z/2} a_s\{r/\kappa(2-r/\kappa)\}^s + O(\kappa^{-(z+2)/2}), \quad 0 \le \theta < \pi/2. \tag{B.1}$$

where $a_s := (2s-2)!!/\{(2s-1)!!s\}$. For $z \ge 2$ of even, the tth power of the both hand of (B.1) is equal to,

$$\begin{aligned}
\theta^{2t} &= \left[\sum_{s=1}^{z/2} a_s\{r/\kappa(2-r/\kappa)\}^s + O(\kappa^{-(z+2)/2})\right]^t \\
&= \sum_{\substack{\sum_{s=1}^{z/2} t_s = t, \\ t \le \sum_s^{z/2} st_s \le z/2}} \frac{t!}{t_1!t_2!\cdots t_{z/2}!} \prod_{l=1}^{z/2} [a_l^{t_l}\{r/\kappa(2-r/\kappa)\}^{lt_l}] + O(\kappa^{-(z+2)/2}) \\
&= \sum_{q=t}^{z/2} A_q(z,t)\{r/\kappa(2-r/\kappa)\}^q + O(\kappa^{-(z+2)/2}), \quad 0 \le \theta < \pi/2. \qquad \text{(B.2)}
\end{aligned}$$

We show that $\int_{\pi/2}^{\pi} K_\kappa(\theta)\theta^j d\theta$ can be ignored.

$$\begin{aligned}
\int_{\pi/2}^{\pi} K_\kappa(\theta)\theta^j d\theta &< \pi^j \int_{\pi/2}^{\pi} K_\kappa(\theta) d\theta \\
&= \pi^j C_\kappa^{-1}(L) \int_\kappa^{2\kappa} L(r)\kappa^{-1/2} r^{-1/2} dr\{2 + O(\kappa^{-1})\}^{-1/2}.
\end{aligned} \tag{B.3}$$

$\int_\kappa^\infty L(r) r^{-1/2} dr = O(\kappa^{-(v+2)/2})$ is given by (v). This leads to,

$$\begin{aligned}
\int_\kappa^{2\kappa} L(r)\kappa^{-1/2} r^{-1/2} dr &= \int_\kappa^\infty L(r) r^{-1/2} dr - \int_{2\kappa}^\infty L(r) r^{-1/2} dr \\
&= O(\kappa^{(-v+2)/2}).
\end{aligned} \tag{B.4}$$

It follows from (1), (B.3) and (B.4) that,

$$\int_{\pi/2}^{\pi} K_\kappa(\theta)\theta^j d\theta = O(\kappa^{(-v+2)/2}). \tag{B.5}$$

By (2), (B.2) and (B.5), the term $\alpha_{2t}(K_\kappa)$ is reduced to,

$$
\begin{aligned}
\alpha_{2t}(K_\kappa) &= 2\int_0^{\pi/2} K_\kappa(\theta)\theta^{2t}d\theta + 2\int_{\pi/2}^{\pi} K_\kappa(\theta)\theta^{2t}d\theta \\
&= 2\int_0^{\pi/2} K_\kappa(\theta)\theta^{2t}d\theta + O(\kappa^{-(v+2)/2}) \\
&= 2C_\kappa^{-1}(L)\int_0^{\kappa} L(r)\sum_{q=t}^{z/2} A_q(z,t)\{r/\kappa(2-r/\kappa)\}^q[r\kappa\{2-r/\kappa\}]^{-1/2}dr + O(\kappa^{-(v+2)/2}) \\
&= 2C_\kappa^{-1}(L)\kappa^{-1/2}\sum_{q=t}^{z/2}\sum_{m=0}^{z/2-q}\frac{A_q(z,t)g_{2q}^{(m)}(0)}{\kappa^{(q+m)}m!}\int_0^{\kappa} L(r)r^{(2q+2m-1)/2}dr + O(\kappa^{-(z+2)/2}) \\
&= 2C_\kappa^{-1}(L)\kappa^{-1/2}\sum_{q=t}^{z/2}\sum_{m=0}^{z/2-q}\frac{A_q(z,t)g_{2q}^{(m)}(0)}{\kappa^{(q+m)}m!}\mu_{2(q+m)}(L) + O(\kappa^{-(z+2)/2}).
\end{aligned}
$$

If K_κ is pth-order kernel, then for $z=p$, the equation (3) is equal to,

$$
\begin{aligned}
\alpha_{2t}(K_\kappa) &= 2C_\kappa^{-1}(L)\kappa^{-1/2}\sum_{q=t}^{p/2}\sum_{m=0}^{p/2-q}\kappa^{-(q+m)}A_q(p,t)(m!)^{-1}g_{2q}^{(m)}(0)\mu_{2(q+m)}(L) + O(\kappa^{-(p+2)/2}) \\
&= 2C_\kappa^{-1}(L)\kappa^{-(p+1)/2}\sum_{q=t}^{p/2}A_q(p,t)(\{p/2-q\}!)^{-1}g_{2q}^{(p/2-q)}(0)\mu_p(L) + O(\kappa^{-(p+2)/2}) \\
&= 2^{1/2}\sum_{q=t}^{p/2}\frac{A_q(p,t)g_{2q}^{(p/2-q)}(0)}{\{p/2-q\}!}\mu_0^{-1}(L)\mu_p(L)\kappa^{-p/2} + O(\kappa^{-(p+2)/2}) \\
&= b_{p,2t}\mu_0^{-1}(L)\mu_p(L)\kappa^{-p/2} + O(\kappa^{-(p+2)/2}).
\end{aligned}
$$

□

Appendix-C

Proof of Lemma 3. If κ is large, then $\delta_{\kappa^{1/2},2t}(L) = \delta_{2t}(L) + o(1)$ since as $\kappa \to \infty$, $\int_{\kappa^{1/2}\pi}^{\infty} L(z^2/2)z^{2t}dz = 0$ and $\delta_{2t}(L)$ is bounded from (iii). Using the Taylor expansion of $\cos(\kappa^{-1/2}z) = 1 - z^2/(2\kappa) + O(\kappa^{-2})$ derives $L(\kappa\{1 -$

$\cos(\kappa^{-1/2}z)\}) = L(z^2/2) + O(\kappa^{-1})$. It is given by these and (2) that,

$$
\begin{aligned}
R(K_\kappa(u)u^t) &= \int_{-\pi}^{\pi} \{K_\kappa(u)u^t\}^2 du \\
&= C_\kappa^{-2}(L) \int_{-\kappa^{1/2}\pi}^{\kappa^{1/2}\pi} L^2(\kappa\{1-\cos(\kappa^{-1/2}z)\})(\kappa^{-1/2}z)^t \kappa^{-1/2} dz \\
&= C_\kappa^{-2}(L)\kappa^{-(2t+1)/2} \int_{-\kappa^{1/2}\pi}^{\kappa^{1/2}\pi} [L(z^2/2) + O(\kappa^{-1})] z^{2t} dz \\
&= C_\kappa^{-2}(L)\kappa^{-(2t+1)/2}[\delta_{2t}(L) + o(1)] \\
&= 2^{-1}\mu_0^{-2}(L)\kappa^{-(2t-1)/2}[\delta_{2t}(L) + o(1)] \\
&= \kappa^{-(2t-1)/2}[d_{2t}(L) + o(1)]. \qquad \text{(C.1)}
\end{aligned}
$$

□

Appendix-D

Proof of Theorem 1. Lemma 2 leads to the expectation:

$$
\begin{aligned}
\mathrm{E}_f[K_\kappa(\theta - Y)] &= \int_{-\pi}^{\pi} K_\kappa(\theta - y) f(y) dy \\
&= \int_{-\pi}^{\pi} K_\kappa(u) \left[\sum_{j=0}^{p+1} \frac{f^{(j)}(\theta)}{j!} u^j + O(u^{p+2}) \right] du \\
&= \sum_{t=0}^{p/2} \frac{f^{(2t)}}{2t!}(\theta)\alpha_{2t}(K_\kappa) + O(\alpha_{p+2}(K_\kappa)) \\
&= f(\theta) + \mu_0^{-1}(L)\mu_p(L)\kappa^{-p/2} \sum_{t=1}^{p/2} \frac{b_{p,2t} f^{(2t)}(\theta)}{2t!} + O(\kappa^{-(p+2)/2}). \qquad \text{(D.1)}
\end{aligned}
$$

It follows from (D.1) that,

$$
\mathrm{bias}[\hat{f}_\kappa(\theta)] = \mu_0^{-1}(L)\mu_p(L)\kappa^{-p/2} \sum_{t=1}^{p/2} \frac{b_{p,2t} f^{(2t)}(\theta)}{2t!} + O(\kappa^{-(p+2)/2}). \qquad \text{(D.2)}
$$

It is given by (4) that,

$$
\begin{aligned}
\mathrm{E}_f[K_\kappa^2(\theta - Y)] &= \int_{-\pi}^{\pi} K_\kappa^2(\theta - y) f(y) dy \\
&= \int_{-\pi}^{\pi} K_\kappa^2(u)[f(\theta) + f'(\theta)u + O(u^2)]du \\
&= R(K_\kappa) f(\theta) + O(R(K_\kappa(u)u)) \\
&= \kappa^{1/2}[f(\theta)d(L) + o(1)]. \qquad \text{(D.3)}
\end{aligned}
$$

By (D.1) and (D.3), the variance is reduced to,

$$
\begin{aligned}
n^{-1}\mathrm{Var}_f[K_\kappa(\theta - Y)] &= n^{-1}\mathrm{E}_f[K_\kappa^2(\theta - Y)] - n^{-1}\mathrm{E}_f[K_\kappa(\theta - Y)]^2 \\
&= n^{-1}[\kappa^{1/2}\{f(\theta)d(L) + o(1)\} - n^{-1}\{f(\theta) + o(1)\}^2] \\
&= n^{-1}\kappa^{1/2} f(\theta) d(L) + o(n^{-1}\kappa^{1/2}). \qquad \text{(D.4)}
\end{aligned}
$$

(5) is derived from (D.2) and (D.4). □

4 Appendix-E

Proof of Theorem 2.

$$
\sqrt{n\kappa^{-1/2}}[\hat{f}_\kappa(\theta) - f(\theta)] = \sqrt{n\kappa^{-1/2}}\{\hat{f}_\kappa(\theta) - \mathrm{E}[\hat{f}_\kappa(\theta)]\} + \sqrt{n\kappa^{-1/2}}\mathrm{bias}[\hat{f}_\kappa(\theta)]. \qquad \text{(E.1)}
$$

Notice that,

$$
\sqrt{n\kappa^{-1/2}}\{\hat{f}_\kappa(\theta) - \mathrm{E}_f[\hat{f}_\kappa(\theta)]\} = \sqrt{n}\left\{ n^{-1}\sum_{i=1}^{n} \kappa^{-1/4} K_\kappa(\theta - \Theta_i) - \mathrm{E}_f[\kappa^{-1/4}K_\kappa(\theta - \Theta_1)] \right\}.
$$

We derive the expectation and the variance of $\kappa^{-1/4}K_\kappa(\theta - \Theta_1)$. It follows from (D.1) that,

$$
\begin{aligned}
\mathrm{E}_f[\kappa^{-1/4}K_\kappa(\theta - \Theta_1)] &= \kappa^{-1/4}\mathrm{E}_f[K_\kappa(\theta - \Theta_1)] \\
&= \kappa^{-1/4}\left[f(\theta) + \mu_0^{-1}(L)\mu_p(L)\kappa^{-p/2}\sum_{t=1}^{p/2}\frac{b_{p,2t}f^{(2t)}(\theta)}{2t!} + O(\kappa^{-(p+2)/2}) \right]. \qquad \text{(E.2)}
\end{aligned}
$$

From (E.2), $|\mathrm{E}[\kappa^{-1/4}K_\kappa(\theta - \Theta_1)]| < \infty$ is derived.

It follows from (D.1) and (D.3) that,

$$\begin{aligned}\mathrm{Var}_f[\kappa^{-1/4}K_\kappa(\theta-\Theta_1)] &= \kappa^{-1/2}\left\{\mathrm{E}_f[K_\kappa^2(\theta-\Theta_1)]-\mathrm{E}_f[K_\kappa(\theta-\Theta_1)]^2\right\}\\ &= \kappa^{-1/2}\left[\kappa^{1/2}\{f(\theta)d(L)+o(1)\}-\{f(\theta)+o(1)\}^2\right]\\ &= f(\theta)d(L)+o(1). \qquad (\mathrm{E.3})\end{aligned}$$

From (E.3), we obtain $\mathrm{Var}_f[\kappa^{-1/4}K_\kappa(\theta-\Theta_1)]<\infty$. From (E.2) and (E.3), the first term of (E.1) satisfies the Lindeberg's condition Feller (1968):

$$\sqrt{n\kappa^{-1/2}}\{\hat{f}_\kappa(\theta)-\mathrm{E}_f[\hat{f}_\kappa(\theta)]\}\xrightarrow{d}\mathrm{N}(0,f(\theta)d(L)). \qquad (\mathrm{E.4})$$

With (D.2) and $\kappa=cn^\alpha$, the rate of the second term of (E.1) is given as,

$$\begin{aligned}\sqrt{n\kappa^{-1/2}}\mathrm{bias}[\hat{f}_\kappa] &= O(\sqrt{n\kappa^{-(2p+1)/2}})\\ &= O(n^{(2-\alpha(2p+1))/4}). \qquad (\mathrm{E.5})\end{aligned}$$

If $\alpha>2/(2p+1)$ is chosen, then the convergent rate of (E.5) is equal to,

$$\sqrt{n\kappa^{-1/2}}\mathrm{bias}[\hat{f}_\kappa]=o(1). \qquad (\mathrm{E.6})$$

For $\alpha>2/(2p+1)$ and as $n\to\infty$, Theorem 2 completes the proof from (E.4) and (E.6). □

Appendix-F

Proof of Theorem 1. It follows from (ii) and (iv) that,

$$\begin{aligned}\mu_j(L_{[p+2]}) &= \int_0^\infty L_{[p+2]}(r)r^{(j-1)/2}dr\\ &= \frac{p+1}{p}\mu_j(L_{[p]})+\frac{2}{p}[L_{[p]}(r)r^{(j+1)/2}]_0^\infty-\frac{2}{p}\frac{j+1}{2}\int_0^\infty L_{[p]}(r)r^{(j-1)/2}dr\\ &= \left(\frac{p+1}{p}-\frac{j+1}{p}\right)\mu_j(L_{[p]}). \qquad (\mathrm{F.1})\end{aligned}$$

Since $K_{[p]}$ is pth -order kernel, the equation (F.1) is reduced to,

$$\mu_0(L_{[p+2]})=\mu_0(L_{[p]})\quad,\quad \mu_j(L_{[p+2]})=0,\quad j=2,4,\ldots,p,\qquad \mu_{p+2}(L_{[p+2]})=-\frac{2}{p}\mu_{p+2}(L_{[p]}).$$

Accodingly, if $L_{[p+2]}$ is (8), $K_{[p+2]}$ has $(p+2)$th-order kernel. □

Appendix-G

Proof of Theorem 2. Put $W := \hat{f}_\kappa(\theta) - \mathrm{I}_\kappa(\theta)$, $\quad Z := \hat{f}_{\kappa/4}(\theta) - \mathrm{I}_{\kappa/4}(\theta)$. Terrell and Scott (1980) showed the follows,

$$\mathrm{E}_f[\hat{f}_\kappa^{\mathrm{TS}}(\theta)] \simeq \mathrm{I}_\kappa(\theta)^{4/3}\mathrm{I}_{\kappa/4}(\theta)^{-1/3}, \tag{G.1}$$

$$\mathrm{Var}_f[\hat{f}_\kappa^{\mathrm{TS}}(\theta)] \simeq \mathrm{Var}_f\left[\frac{4}{3}W - \frac{1}{3}Z\right]. \tag{G.2}$$

When K_κ is second-order kernel and z=4, the equation (3) is equal to,

$$\alpha_{2t}(K_\kappa) = 2^{1/2}\mu_0^{-1}(L)\sum_{q=t}^{3}\sum_{m=0}^{3-q}\frac{A_q(4,t)g_{2q}^{(m)}(0)\mu_{2(q+m)}(L)}{\kappa^{q+m}m!} + O(\kappa^{-3}). \tag{G.3}$$

Let be,

$$\begin{aligned}
\alpha_{2,1} &:= 2^{1/2}\mu_0^{-1}(L)\mu_2(L)A_1(4,1)g_2(0),\\
\alpha_{2,2} &:= 2^{1/2}\mu_0^{-1}(L)\mu_4(L)[A_1(4,1)g_2^{(1)}(0) + A_2(4,1)g_4(0)],\\
\alpha_{4,1} &:= 2^{1/2}\mu_0^{-1}(L)\mu_4(L)A_2(4,2)g_4(0).
\end{aligned}$$

Then, we can write (G.3) as,

$$\begin{aligned}
\alpha_2(K_\kappa) &= \alpha_{2,1}\kappa^{-1} + \alpha_{2,2}\kappa^{-2} + O(\kappa^{-3}),\\
\alpha_4(K_\kappa) &= \alpha_{4,1}\kappa^{-2} + O(\kappa^{-3}),\\
\alpha_6(K_\kappa) &= O(\kappa^{-3}).
\end{aligned}$$

Let be $c_j := f^{(j)}/j!$ and $I_\kappa(\theta) := \mathrm{E}_f[\hat{f}_\kappa(\theta)]$. By the simular procedure as Terrell and Scott (1980), $\log I_\kappa(\theta)$ is given as,

$$\log I_\kappa(\theta) = \log f(\theta) + \frac{c_2\alpha_{2,1}}{f(\theta)\kappa} + \frac{(c_2\alpha_{2,2} + c_4\alpha_{4,1})f(\theta) - c_2^2\alpha_{2,1}^2/2}{f^2(\theta)\kappa^2} + O(\kappa^{-3}). \tag{G.4}$$

For (G.4), taking exponentials of $\{4\log I_\kappa(\theta)/3 - \log I_{\kappa/4}(\theta)/3\}$ gives the follows,

$$I_\kappa(\theta)^{4/3}I_{\kappa/4}(\theta)^{-1/3} = f(\theta) + 4\frac{c_2^2\alpha_{2,1}^2/2 - (c_2\alpha_{2,2} + c_4\alpha_{4,1})f(\theta)}{f(\theta)\kappa^2} + O(\kappa^{-3}). \tag{G.5}$$

It follows from (G.1) and (G.5) that,

$$\text{bias}_f[\hat{f}_\kappa^{\text{TS}}(\theta)] = G(\theta)\kappa^{-2} + O(\kappa^{-3}), \tag{G.6}$$

where,

$$G(\theta) := 4\frac{c_2^2\alpha_{2,1}^2/2 - (c_2\alpha_{2,2} + c_4\alpha_{4,1})f(\theta)}{f(\theta)}. \tag{G.7}$$

Let be $\delta_{\kappa^{1/2},2t,4}(L) := \int_{-\kappa^{1/2}\pi}^{\kappa^{1/2}\pi} L(z^2/2)L(z^2/8)z^{2t}dz$ and $\delta_{2t}(L) := \int_{-\infty}^{\infty} L(z^2/2)L(z^2/8)z^{2t}dz$. Then, $\delta_{2t,4}(L) < \infty$ and $\delta_{\kappa^{1/2},2t,4}(L) = \delta_{2t,4}(L) + o(1)$ since for all z, $L(z^2/2) > L(z^2/8)$ and $\delta_{2t}(L) < \infty$. In the same way as (4), this leads to,

$$\begin{aligned}\int_{-\pi}^{\pi} K_\kappa(u)K_{\kappa/4}(u)u^{2t}du &= C_\kappa^{-1}(L)C_{\kappa/4}^{-1}(L)\int_{-\kappa^{1/2}\pi}^{\kappa^{1/2}\pi} L_\kappa(\kappa^{-1/2}z)L_{\kappa/4}(\kappa^{-1/2}z)(\kappa^{-1/2}z)^{2t}\kappa^{-1/2}dz \\ &= \kappa^{-(2t-1)/2}[d_{2t,4}(L) + o(1)],\end{aligned} \tag{G.8}$$

where $d_{2t,4}(L) := 2^{-2}\mu_0^{-2}(L)\delta_{2t,4}(L)$. (G.8) gives the follows,

$$\begin{aligned}\text{E}_f[K_\kappa(\theta - Y)K_{\kappa/4}(\theta - Y)] &= \int_{-\pi}^{\pi} K_\kappa(\theta - y)K_{\kappa/4}(\theta - y)f(\theta)dy \\ &= \int_{-\pi}^{\pi} K_\kappa(u)K_{\kappa/4}(u)duf(\theta) + O\left(\int_{-\pi}^{\pi} K_\kappa(u)K_{\kappa/4}(u)u^2du\right) \\ &= \kappa^{1/2}[f(\theta)d_{0,4}(L) + o(1)].\end{aligned} \tag{G.9}$$

It is given by (D.1) and (G.9) that,

$$\begin{aligned}\text{Cov}_f[WZ] &= n^{-1}\text{Cov}[K_\kappa(\theta - Y), K_{\kappa/4}(\theta - Y)] \\ &= n^{-1}[\kappa^{1/2}\{f(\theta)d_{0,4}(L) + o(1)\} - \{f(\theta) + o(1)\}^2] \\ &= n^{-1}\kappa^{1/2}f(\theta)d_{0,4}(L) + o(n^{-1}\kappa^{1/2}).\end{aligned} \tag{G.10}$$

By (D.4) and (G.10), the equation (G.2) is reduced to,

$$\begin{aligned}\text{Var}_f[\hat{f}_\kappa^{\text{TS}}(\theta)] &\simeq \text{Var}_f\left[\frac{4}{3}W - \frac{1}{3}Z\right] \\ &= \frac{16}{9}\text{Var}_f[W] - \frac{8}{9}\text{Cov}_f[ZW] + \frac{1}{9}\text{Var}_f[Z^2] \\ &= n^{-1}\kappa^{1/2}f(\theta)D(L) + o(n^{-1}\kappa^{1/2}),\end{aligned} \tag{G.11}$$

where,

$$D(L) := \frac{33d(L) - 16d_{0,4}(L)}{18}. \tag{G.12}$$

□

Appendix-H

Recall that $f_{\mathrm{VM}}(\theta;\tau)$ is von Mises density. Then,

$$\hat{R}_{\hat{\tau}}(f_{\mathrm{VM}}^{(2)}) = \frac{3\hat{\tau}^2 I_2(2\hat{\tau}) + 2\hat{\tau} I_1(2\hat{\tau})}{8\pi I_0^2(\hat{\tau})},$$

$$\hat{R}_{\hat{\tau}}\left(\frac{5f_{\mathrm{VM}}^{(2)}}{12} + \frac{f_{\mathrm{VM}}^{(4)}}{6}\right) = \frac{25\hat{R}_{\hat{\tau}}(f_{\mathrm{VM}}^{(2)})}{144} - \frac{5\hat{R}_{\hat{\tau}}(f_{\mathrm{VM}}^{(3)})}{36} + \frac{\hat{R}_{\hat{\tau}}(f_{\mathrm{VM}}^{(4)})}{36},$$

where,

$$\hat{R}_{\hat{\tau}}(f_{\mathrm{VM}}^{(2)}) = \frac{3\hat{\tau}^2 I_2(2\hat{\tau}) + 2\hat{\tau} I_1(2\hat{\tau})}{8\pi I_0^2(\hat{\tau})},$$

$$\hat{R}_{\hat{\tau}}(f_{\mathrm{VM}}^{(3)}) = \frac{4\hat{\tau} I_1(2\hat{\tau}) + 30\hat{\tau}^2 I_2(2\hat{\tau}) + 15\hat{\tau}^3 I_3(2\hat{\tau})}{16\pi I_0^2(\hat{\tau})},$$

$$\hat{R}_{\hat{\tau}}(f_{\mathrm{VM}}^{(4)}) = \frac{8\hat{\tau}^2 I_0(2\hat{\tau}) + 105\hat{\tau}^4 I_2(2\hat{\tau}) + 105\hat{\tau}^3 I_3(2\hat{\tau}) + 244\hat{\tau}^2 I_2(2\hat{\tau})}{32\pi I_0^2(\hat{\tau})}.$$

$$\begin{aligned}\hat{R}_{\hat{\tau}}\left(\frac{\{f_{\mathrm{VM}}^{(2)}\}^2}{2f_{\mathrm{VM}}} - \frac{5f_{\mathrm{VM}}^{(2)} + 2f_{\mathrm{VM}}^{(4)}}{4}\right) &= \frac{1}{4}\hat{R}_{\hat{\tau}}(f_{\mathrm{VM}}^{(2)}/f_{\mathrm{VM}}^{-1}) + \frac{25}{16}\hat{R}_{\hat{\tau}}(f_{\mathrm{VM}}^{(2)}) \\ &- \frac{5}{4}\hat{R}_{\hat{\tau}}(f_{\mathrm{VM}}^{(3)}) + \frac{1}{4}\hat{R}_{\hat{\tau}}(f_{\mathrm{VM}}^{(4)}) \\ &- \frac{5}{4}\int_{-\pi}^{\pi}\{f_{\mathrm{VM}}^{(2)}(\theta)\}^3 f_{\mathrm{VM}}^{-1}(\theta)d\theta \\ &- \frac{1}{2}\int_{-\pi}^{\pi}\{f_{\mathrm{VM}}^{(2)}(\theta)\}^2 f_{\mathrm{VM}}^{(4)}(\theta) f_{\mathrm{VM}}^{-1}(\theta)d\theta,\end{aligned}$$

where,

$$\hat{R}_{\hat{\tau}}(f_{\mathrm{VM}}^{(2)}/f_{\mathrm{VM}}^{-1}) = \frac{41\hat{\tau}^4 I_2(2\hat{\tau}) + 12\hat{\tau}^2 I_2(\hat{\tau}) - 87\hat{\tau}^3 I_3(2\hat{\tau})}{32\pi I_0(\hat{\tau})^2},$$

$$\int_{-\pi}^{\pi}\{f^{(2)}(\theta)\}^3 f^{-1}(\theta)d\theta = \frac{4\hat{\tau}^3 I_1(2\hat{\tau}) - 14\hat{\tau}^2 I_2(2\hat{\tau}) - 3\hat{\tau}^3 I_3(2\hat{\tau})}{16\pi I_0(\hat{\tau})^2},$$

$$\int_{-\pi}^{\pi}\{f_{\mathrm{VM}}^{(2)}(\theta)\}^2 f_{\mathrm{VM}}^{(4)}(\theta) f_{\mathrm{VM}}^{-1}(\theta)d\theta = \frac{36\hat{\tau}^2 I_2(2\hat{\tau}) + 9\hat{\tau}^4 I_2(2\hat{\tau}) + 25\hat{\tau}^3 I_3(2\hat{\tau})}{32\pi I_0(\hat{\tau})^2}.$$