Annals of the Institute of Statistical Mathematics (2023) 75:855–886
https://doi.org/10.1007/s10463-023-00868-4

Automatic data-based bin width selection for rose diagram

Yasuhito Tsuruta[1] · **Masahiko Sagae**[2]

Received: 9 May 2022 / Revised: 18 January 2023 / Accepted: 24 January 2023 /
Published online: 9 March 2023

Abstract
A rose diagram is a representation that circularly organizes data with the bin width as the central angle. This diagram is widely used to display and summarize circular data. Some studies have proposed the selector of bin width based on data. However, only a few papers have discussed the property of these selectors from a statistical perspective. Thus, this study aims to provide a data-based bin width selector for rose diagrams using a statistical approach. We consider that the radius of the rose diagram is a nonparametric estimator of the square root of two times the circular density. We derive the mean integrated square error of the rose diagram and its optimal bin width and propose two new selectors: normal reference rule and biased cross-validation. We show that biased cross-validation converges to its optimizer. Additionally, we propose a polygon rose diagram to enhance the rose diagram.

Keywords Rose diagram · Bin width estimator · Circular data · Nonparametric density estimator · Histogram estimator

1 Introduction

Suppose that circular data $\theta_1, \ldots, \theta_n$ are a random sample from a circular density $f(\theta)$ for $\theta \in [0, 2\pi)$. Rose diagrams are frequently used to display and summarize circular data, such as wind orientation, paleocurrent direction, and bird migration direction. A rose diagram circularly arranges sectors with the bin width as the central angle (see Fig. 1). Some researchers have asserted that the ancestor of modern rose diagrams was Nightingale (1858), who used it to show monthly variations in military mortalities (Fisher 1993; Chapter 1; Wells 2000). Sanderson and Peacock

✉ Yasuhito Tsuruta
tsuruta.yasuhito@u-nagano.ac.jp

[1] Faculty of Global Management Studies, The University of Nagano, 8-49-7, Miwa, Nagano, Nagano 380-8525, Japan

[2] School of Economics, Kanazawa University, Kakumamachi, Kanazawa, Ishikawa 920-1192, Japan

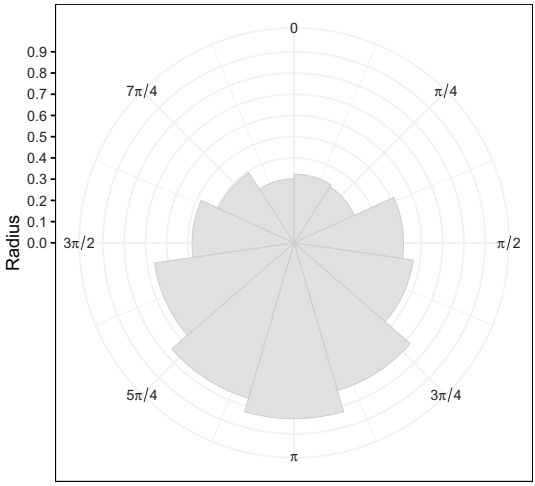


Fig. 1 Rose diagram summarizing the sample ($n = 500$) being generated from its von Mises distribution (See Fig. 2) with bin width $h = 2\pi/11$ (being estimated based on the equal normal reference rule)

(2020) found that rose diagrams are used to (i) show a departure from a random or uniform distribution of direction, (ii) establish the poly-modal nature of data, (iii) compare relative strengths of different modes, and (iv) compare regional variations in the data.

Many studies have recommended that the area of a sector should be proportional to the frequency by making the radius proportional to the square root of the frequency (Mardia 1972, Chapter 1; Nemec 1988; Fisher 1993, Chapter 2; Mardia and Jupp 2000, Chapter 1; Sanderson and Peacock 2020). This is because, when the radius of a sector is proportional to its frequency, the area of that sector is also proportional to the square of the frequency, causing the peak to appear exaggerated (Nemec 1988). Therefore, we employ the radius proportional to the square root of the relative frequency, such as the total area of the sectors is one, as proposed by Nemec (1988).

The parameters controlling a rose diagram are the starting point of the bins and the bin width. We assume that the starting point t_1 is zero without losing generality, as the origin of the circular data is chosen arbitrarily. For definitions of the m bins, we follow mesh $\{t_l\}, \quad l = 1, 2, \ldots, m, m+1$ with $t_{m+1} = 2\pi$ and $t_1 < t_2 < \cdots < t_m < t_{m+1}$. Define k-th bin as $B_k := [t_k, t_{k+1})$ with its bin width $h_k := |t_{k+1} - t_k|$ and k-th bin's frequency as v_k for $k = 1, 2, \ldots, m$. Set $\lceil x \rceil := \min\{n \in \mathbb{Z} | x \le n\}$. We insert that the first $m-1$ bin widths h_k for $k = 1, \ldots, m-1$ are a common length h, and the m-th bin width is only $h_m = ch$ with any constant $c \in (0, 1]$, where $c = 2\pi/h - (m-1)$ with $m := \lceil 2\pi/h \rceil$. If $2\pi/h$ is not an integer, then, c is its fractional part. We then define the radius of the rose diagram as follows:

$$\hat{r}_f(\theta; h) := \sqrt{2v_k/(nh_k)} = \begin{cases} \sqrt{2v_k/(nh)} & \text{if } \theta \in B_k, \quad k = 1, 2, \ldots, m-1, \\ \sqrt{2v_k/(nch)} & \text{if } \theta \in B_m. \end{cases} \tag{1}$$

From Eq. (1), we find that selection of the common bin width h has an important role in controlling the shape of a rose diagram. In this regard, studies have also discussed

the choice of bin width. Andreasen (1990) suggested that the bin width should be equal to the confidence interval width of the sample circular mean. Other studies have proposed that the sample size should determine the bin width. For example, Ballantyne and Cornish (1979) claimed that the bin number should not exceed $n/5$. Pewsey et al. (2013) contended that the square root of the sample size is often used as a reasonable first guess at the appropriate number of bins. Meanwhile, Sanderson and Peacock (2020) proposed a bin width selection rule $h = R/(2n^{1/3})$, where R is the range 2π. The basis of $h \propto n^{-1/3}$ is the result of Scott (1979) for a histogram on the line. Scott (1979) provided the mean integrated square error (MISE) between a histogram estimator and an unknown density and derived the optimal bin width as its minimizer. To the best of our knowledge, in these rules, consistency has not been established thus far.

This study aims to provide an automatic data-based bin width selection rule for a rose diagram with consistency. First, we discuss the relationship between the rose diagram and an unknown circular density. We present a circular distribution function $F(\theta)$ as a sector, wherein the radius function $r_f(\theta)$, subtracted from the origin, moves as much as θ, as shown in Fig. 2. Then, note that:

$$r_f(\theta) := \sqrt{2f(\theta)}. \tag{2}$$

By comparing (1) and (2), we find that the radius of the rose diagram is the estimator of $r_f(\theta)$. For the global error criteria between $\hat{r}_f(\theta;h)$ and $r_f(\theta)$, we employ the integrated square error (ISE) $\mathrm{ISE}[\hat{r}_f(\theta;h)] := \int_0^{2\pi} \{\hat{r}_f(\theta;h) - r_f(\theta)\}^2 d\theta$ and its MISE; that is, $\mathrm{MISE}[\hat{r}_f(\theta;h)] := \mathrm{E}[\mathrm{ISE}[\hat{r}_f(\theta;h)]]$. We derive the asymptotic MISE and optimizer $h_{\mathrm{opt}} = O(n^{-1/3})$ and, thus, obtain the convergence rate of the optimal MISE as $O(n^{-2/3})$. Through these means, we propose two new bin width selection rules for rose diagrams: first, the normal reference rule (NRR), which is obtained under the assumption that f is the von Mises density, and second, the biased cross-validation (BCV). We show that this is consistent with the optimal bin width.

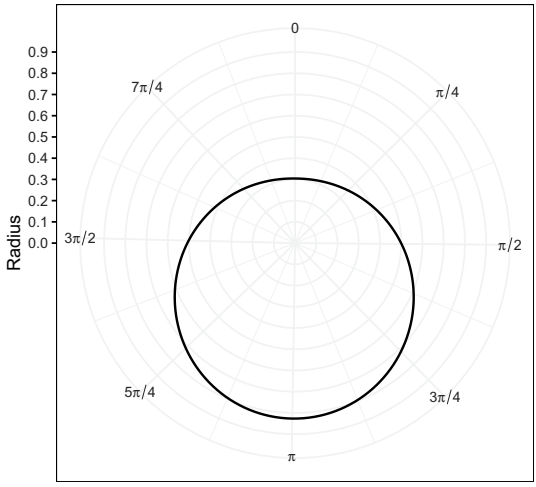


Fig. 2 Von Mises distribution with mean $\mu = \pi$ and concentration parameter $\kappa = 1$. This radius is $\sqrt{2f(\theta)}$ with density f

However, a circular kernel density estimator shows that the convergence rate of the minimizing MISE is $O(n^{-4/5})$ (Hall et al. 1987; Taylor 2008; Di Marzio et al. 2011; Tsuruta and Sagae 2017, 2020). The rose diagram is less accurate than the circular kernel density estimator because the bins are not smooth. We apply the linear interpolation to the rose diagram to overcome this lack of smoothness; then, we call the corresponding estimator the polygon rose diagram estimator, based on the frequency polygon proposed by Scott (1985). We show that the polygon rose diagram estimator has a convergence rate of $O(n^{-4/5})$, which is equivalent to the kernel density estimation.

The remainder of this study is organized as follows: In Sect. 2, we discuss the optimal bin width of the rose diagram and its bin width selectors. In Sect. 3, we provide the properties of the polygon rose diagram estimator. Section 4 summarizes the results of the numerical experiment. Section 5 illustrates the real-data example for these estimators. Finally, Sect. 6 concludes.

2 Optimal bin width of the rose diagram estimator

This section is structured as follows: We derive the optimal bin width of the rose diagram estimator and describe NRR in Sect. 2.1. Section 2.2 summarizes the asymptotic properties of BCV. For simplicity, we admit that the periodic extension of f to $\mathbb{R}$ given by $f(\theta + 2\pi) = f(\theta)$ for any $\theta \in \mathbb{R}$ such as $\int_a^{a+2\pi} f(\theta)d\theta = 1$ for any $a \in \mathbb{R}$.

2.1 Optimal width under von Mises density

First, we investigate the theoretical properties of the rose diagram estimator $\hat{r}_f(\theta;h)$. Let $R(g) := \int_0^{2\pi} g(\theta)^2 d\theta$. We set four more assumptions:

(A1) Bin width $h =: h_n \to 0$ and $nh \to \infty$, as $n \to \infty$.
(A2) For any θ, $f(\theta) > 0$.
(A3) Density f is the second continuously differentiable.
(A4) First derivative $r_f'(\theta)$ is bounded for any θ. Quantity $R(r_f')$ is bounded.

Assumptions (A1) and (A4) are required to ensure the asymptotic MISE goes to zero as n increases. Noting that $r_f'(\theta) = f'(\theta)/[\sqrt{2f(\theta)}]$ and $r_f''(\theta) = f''(\theta)/[\sqrt{2f(\theta)}] - f'(\theta)^2/[2\sqrt{2}f(\theta)^{3/2}]$, assumptions (A2) and (A3) imply that $r_f(\theta)$ is the second differentiable.

We recognize that the bin count $\{v_k\}$ follows binomial distribution $B(n, p_k)$, where $p_k := \int_{B_k} f(\theta)d\theta$. We know that the k-th bin's relative frequency v_k/n approximates the normal distribution $N(p_k, p_k(1-p_k)/n)$. The mean value theorem indicates that $p_k = \int_{B_k} f(\theta)d\theta = h_k f(\xi_k)$ for some $\xi_k \in B_k$, where $h_k = h$ for $k = 1, \ldots, m-1$ and

$h_m = ch$ with $c \in (0, 1]$. Therefore, we find that the histogram estimator $\hat{f}(\theta;h_k) := v_k/(nh_k)$ for $\theta \in B_k$ has an asymptotic distribution in the following lemma:

Lemma 1 *Let* $W_k := \hat{f}(\theta;h) - \mathrm{E}[\hat{f}(\theta;h)]$. *Assume that* (A1), (A2), (A3), *and* (A4) *hold. Then*,

$$(nh_k)^{1/2} W_k \xrightarrow{d} \mathrm{N}(0, f(\xi_k)),\ \theta \in B_k,$$

as $n \to \infty$. *The moments are* $E[W_k^{2t-1}] = 0$, $E[W_k^2] = f(\xi_k)/(nh_k) - f(\xi_k)^2/n$, *and* $E[W_k^{2t}] = O((nh)^{-t})$ *for a natural number* t.

Let $\mathrm{Bias}[\hat{r}_f(\theta;h)] := \mathrm{E}[\hat{r}_f(\theta;h)] - r_f(\theta)$. By applying the delta method to Lemma 1, we obtain the following asymptotic MISE of the rose diagram estimator:

Theorem 1 *Assume that* (A1), (A2), (A3), *and* (A4) *hold. Then*,

$$\mathrm{Bias}[\hat{r}_f(\theta;h)] = \begin{cases} r_f'(\theta)(h/2 + t_k - \theta) + O(h^2 + (nh)^{-1}) & \text{if } \theta \in B_k, k = 1, \ldots, m-1, \\ r_f'(\theta)[(ch)/2 + t_k - \theta] + O(h^2 + (nh)^{-1}) & \text{if } \theta \in B_m, \end{cases} \tag{3}$$

and

$$\mathrm{Var}[\hat{r}_f(\theta;h)] = \begin{cases} 1/(2nh) + o((nh)^{-1}) & \text{if } \theta \in B_k, k = 1, \ldots, m-1, \\ 1/(2cnh) + o((nh)^{-1}) & \text{if } \theta \in B_m. \end{cases} \tag{4}$$

Combining (3) *and* (4) *leads to*:

$$\mathrm{MISE}[\hat{r}_f(\theta;h)] = \mathrm{AMISE}[\hat{r}_f(\theta;h)] + o(h^2 + (nh)^{-1}),$$

where

$$\mathrm{AMISE}[\hat{r}_f(\theta;h)] = h^2 R(r_f')/12 + \pi/(nh). \tag{5}$$

The minimizer of the right-hand side (*RHS*) *in* (5) *is*:

$$h_{\mathrm{opt}} = [6\pi/R(r_f')]^{1/3} n^{-1/3}. \tag{6}$$

By combining (5) *and* (6), *we find that the convergence rate of the optimal MISE is* $O(n^{-2/3})$.

We present the proof in Appendix A. We derive the corollary from Lemma 1 and Theorem 1 using the delta method.

Corollary 1 *Assume that* (A1), (A2), (A3), *and* (A4) *hold. Then*,

$$(nh)^{1/2}[\hat{r}_f(\theta;h) - r_f(\theta)] \xrightarrow{d} \mathrm{N}((nh)^{1/2}\mathrm{Bias}[\hat{r}_f(\theta;h)], (2c_k)^{-1}), \quad \theta \in B_k, \tag{7}$$

as $n \to \infty$, *where* $c_k := 1$ *for* $k = 1, 2, \ldots, m-1$ *and* $c_m := c$. *Let* $h = n^{\alpha}$. *When* $\alpha < -1/3$ *and* $n \to \infty$, *we eliminate* $(nh)^{1/2}\mathrm{Bias}[\hat{r}_f(\theta;h)]$ *in the RHS of* (7).

In Theorem 1, the rose diagram estimator indicates that the optimal MISE is $O(n^{-2/3})$. This rate is the same as that of the histogram estimator. The variance of the rose diagram estimator does not depend on density f. Each bin count is approximately a Poisson random variable in the histogram, and the square root of the histogram is the variance-stabilizing transformation for Poisson data (Scott 2015, Chapter 3). It is important to note that, according to (Scott 2015, Chapter 3), the square root scale allows for easy comparisons of noise in the histogram in regions of high and low density. Moreover, if selecting all equal bin widths with $c = 1$, then, the rose diagram estimator has the homogeneity of variance.

We find that the asymptotic MISE of the rose diagram relates to the asymptotic mean Hellinger distance (AMHD) of the histogram provided by Kanazawa (1993). The AMHD of a histogram with the support of length 2π is equal to the AMISE of the rose diagram divided by two. However, Kanazawa (1993) assumed that f is the third differentiable. This constriction is more severe than the construction in Theorem 1.

The optimal bin width depends on the sample size and the unknown quantity $R(r_f')$, which denotes the magnitude of the roughness of the first derivative of the underlying density f. Therefore, the bin width should be decreased as the sample size increases with $h \propto n^{-1/3}$. It is preferable to decrease the bin width as the magnitude of the roughness of f increases.

The optimal bin width requires the calculation of $R(r_f')$. We further suggest employing the von Mises density for the reference density, as it plays a central role in the statistical inference for circular data. It is also called the circular normal density. The von Mises density has: $R(r_f') = \kappa I_1(\kappa)/[2I_0(\kappa)]$, where $I_p(\kappa)$ denotes the modified Bessel function of the first kind and order p with its concentration parameter κ. In the literature, Mardia and Jupp (2000, Chapters 3 and 5) explained the statistical properties of the von Mises density, such as its maximum likelihood estimator (MLE) and $I_p(\kappa)$. Along these lines, we propose NRR as follows:

Definition 1 NRR is:

$$\hat{h}_{\mathrm{NRR}} = [12\pi I_0(\hat{\kappa})/\{\hat{\kappa} I_1(\hat{\kappa})\}]^{1/3} n^{-1/3},$$

where $\hat{\kappa}$ is the MLE of κ.

2.2 Optimal nonparametric bin width

We consider a new nonparametric approach for estimating the optimal bin width. This idea is based on the BCV of a density estimator (Scott and Terrell 1987). We approximate $r_f'(t_k)$ by applying a finite difference in the rose diagram at the midpoints of bin B_k and B_{k-1} using the following equation:

$$\tilde{r}'_f(t_k) = \frac{\hat{r}_k - \hat{r}_{k-1}}{\tilde{h}_k}, \tag{8}$$

where $\tilde{h}_k = (h_k + h_{k-1})/2$, $\hat{r}_k := \hat{r}_f(t_k + h_k/2;h) = \sqrt{2v_k/(nh_k)}$ and $\hat{r}_{k-1} := \hat{r}_f(t_k - h_{k-1}/2;h) = \sqrt{2v_{k-1}/(nh_{k-1})}$ for $k = 1, 2, \cdots, m$ with $B_0 := [t_m, t_{m+1}) = B_m$. Equation (8) produces the Riemann approximation of $R(r'_f)$ as follows:

$$\hat{R}_1 = \sum_k \tilde{r}'_f(t_k)^2 \tilde{h}_k = \sum_k \tilde{h}_k^{-1} (\hat{r}_k - \hat{r}_{k-1})^2.$$

However, the expectation is:

$$\mathrm{E}[\hat{R}_1] = R(r'_f) + \frac{2\pi}{nh^3} + o(1 + (nh^3)^{-1}). \tag{9}$$

We present the detail of Eq. (9) in Appendix C. Selecting the optimal h in Eq. (6) indicates that the second term in the RHS of Eq. (9) approximates to $2\pi/(nh^3_{\mathrm{opt}}) = R(r'_f)/3$. Therefore, $\hat{R}_1$ has a bias that is one-third larger than $R(r'_f)$. Correcting this bias yields the following asymptotic unbiased estimator:

$$\hat{R}(r'_f) := \sum_k \frac{(\hat{r}_k - \hat{r}_{k-1})^2}{\tilde{h}_k} - \frac{2\pi}{nh^3}. \tag{10}$$

In fact, we have $\mathrm{E}[\hat{R}(r'_f)] = R(r'_f) + o(1)$. By estimating $R(r'_f)$ using (10) in the AMISE expression (5), we obtain the following function:

$$\mathrm{BCV}(h) = \frac{h^2}{12} \sum_k \frac{(\hat{r}_k - \hat{r}_{k-1})^2}{\tilde{h}_k} + \frac{5\pi}{6nh}. \tag{11}$$

We define the BCV estimator in the following manner:

Definition 2 The BCV estimator $\hat{h}_{\mathrm{BCV}}$ is given by:

$$\hat{h}_{\mathrm{BCV}} := \underset{h \in (0,a)}{\arg\min}\, \mathrm{BCV}(h),$$

for $0 < a \leq 2\pi$.

We recommend selecting $a = \pi$ in practical data analysis because the BCV function often has the local minimum at a value slightly greater than π, although the optimal bin width becomes smaller as n becomes larger (see Fig. 3). In addition, we consider that $a \geq \pi$ is too large to fully visualize the shape of an underlying density, as this choice provides only two bins.

The BCV function has the expectation and variance that follows the theorem mentioned below.

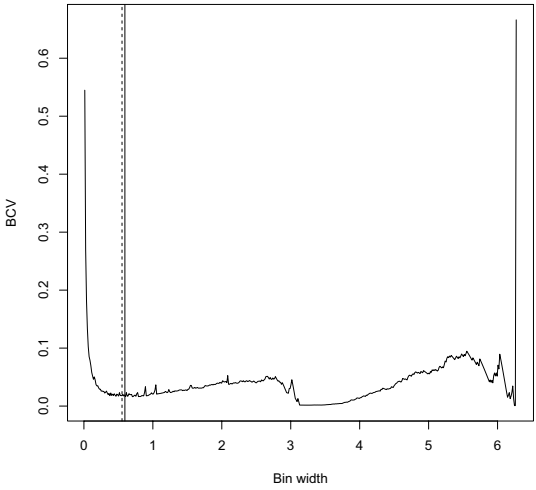


Fig. 3 BCV function in Eq. (11) based on the sample ($n = 500$) generated from the von Mises density with $\mu = \pi$ and $\kappa = 1$. The dashed vertical line is the optimal bin width ($h = 0.553$). The solid vertical line is the value estimated using the BCV estimator with $a = \pi$ ($h = 0.594$)

Theorem 2 *Assume that* (*A*1), (*A*2), (*A*3), *and* (*A*4) *hold. Then*,

$$\mathrm{E}[\mathrm{BCV}(h)] = \mathrm{AMISE}[\hat{r}_f(\theta;h)] + o(h^2 + (nh)^{-1}),$$

and

$$\mathrm{Var}[\mathrm{BCV}(h)] = \frac{1}{18n} + o(n^{-1}).$$

We present the proof in Appendix C. Theorem 2 leads to BCV estimator $\hat{h}_{\mathrm{BCV}}$ being consistent with h_{opt}.

Corollary 2 *Assume that* (*A*1), (*A*2), (*A*3), *and* (*A*4) *hold. Then*,

$$\hat{h}_{\mathrm{BCV}}/h_{\mathrm{opt}} \xrightarrow{p} 1,$$

as $n \to \infty$.

We present the proof in Appendix D.

NRR or BCV often provides $c \neq 1$ for the m-th bin width. The circular support of density f allows statisticians to subjectively determine where the m-th bin is set. To avoid this problem, we consider that all bins should have the same bin width. In this case, we employ $h_m = 2\pi/m$ with a natural number m. A natural number M exists such that it satisfies:

$$h_{M+1} < h_{\mathrm{opt}} < h_M, \tag{12}$$

for the optimal bin width, h_{opt}. Theorem 1 implies that:

$$\frac{\mathrm{AMISE}[\hat{r}_f(\theta;\delta h_{\mathrm{opt}})]}{\mathrm{AMISE}[\hat{r}_f(\theta;h_{\mathrm{opt}})]} = \frac{\delta^2}{3} + \frac{2}{3\delta}, \tag{13}$$

where $\delta := h/h_{\mathrm{opt}}$. The ratio in Eq. (13) has only local minimum at $\delta = 1$. Therefore, if n is sufficiently large, then, the bin widths h_M or h_{M+1} have the smallest of MISE in $\{h_m\}$. Thus, replacing the optimizer h_{opt} in Eq. (12) with any estimator $\hat{h}$ provides an equal split bin width estimator as follows:

Algorithm 1 Let $M := \max\{n \in \mathbb{Z} | n \le 2\pi/\hat{h}\}$ for any bin width estimator $\hat{h}$.

Step 1 Provide $\hat{h}_M = 2\pi/M$ and $\hat{h}_{M+1} = 2\pi/(M+1)$.
Step 2 Provide equal split bin width estimator:

$$\hat{h}_E := \underset{M,M+1}{\arg\min}\{\mathrm{AMISE}[\hat{r}_f(\theta;\hat{h}_M)], \mathrm{AMISE}[\hat{r}_f(\theta;\hat{h}_{M+1})]\}.$$

If they are both the same, then, $\hat{h}_E = \hat{h}_M$.

Algorithm 1 provides an equal split bin width estimator based on NRR by employing the AMISE that estimates $R(r_f')$ in the RHS in Eq. (5) by $\hat{\kappa} I_1(\hat{\kappa})/[2I_0(\hat{\kappa})]$ with the MLE $\hat{\kappa}$. We call this the equal split normal reference rule (ENRR). This also provides the equal split bin biased cross-validation (EBCV) estimator by using AMISE that estimate $R(r_f')$ by Eq. (10).

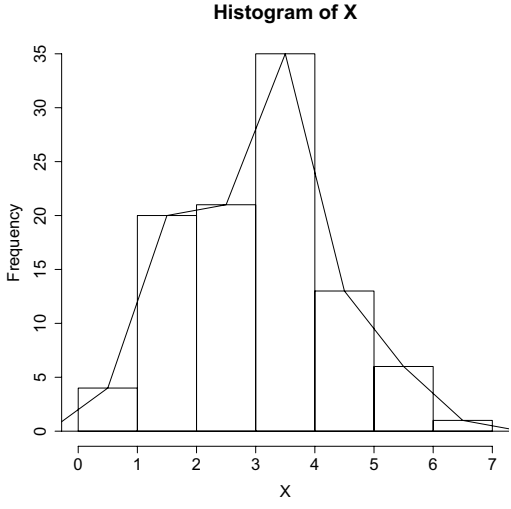


Fig. 4 Frequency polygon is the linear interpolation of the midpoints of a bar in the histogram

3 Improving the rose diagram estimator

The rose diagram estimator has a convergence rate of $O(n^{-2/3})$ in its MISE. However, the circular kernel density estimator has a convergence rate of $O(n^{-4/5})$. This is because the rose diagram is not smooth and has bumps on each bin. Thus, smoothing may improve performance. For linear data, Scott (1985) proposed that the frequency polygon is a continuous density estimator based on a histogram (see Fig. 4). We propose a polygon rose diagram estimator with the linear interpolation of the rose diagram (see Fig. 5):

Definition 3

$$\tilde{r}_f(\theta;h) = \left(\frac{h_k}{2\tilde{h}_k} - \frac{\theta - t_k}{\tilde{h}_k}\right)\hat{r}_{k-1}(\theta;h) + \left(\frac{h_{k-1}}{2\tilde{h}_k} + \frac{\theta - t_k}{\tilde{h}_k}\right)\hat{r}_k(\theta;h),$$

for $\theta \in [t_k - h_{k-1}/2, t_k + h_k/2)$ with $k = 1, 2, \ldots m$, where $\tilde{h}_k$ is the length of each interval and $\hat{r}_k(\theta;h)$ is the rose diagram estimator in the k-th bin B_k with $B_0 = B_m$. Further, the interval $[t_1 - h_m/2, t_1)$ corresponds to the interval $[t_{m+1} - h_m/2, t_{m+1})$ due to the circular support.

This estimator requires a smoother condition of an underlying density to derive its theoretical properties. The assumption is

(A5). Density f is third continuously differentiable. The second Derivative $r_f''(\theta)$ is bounded for any θ. Quantities $R(r_f'')$, $R((r_f'')^{1/2} r_f' r_f^{-1/2})$, and $R((r_f')^2 r_f^{-1}))$ are bounded.

We have its asymptotic MISE as the following theorem.

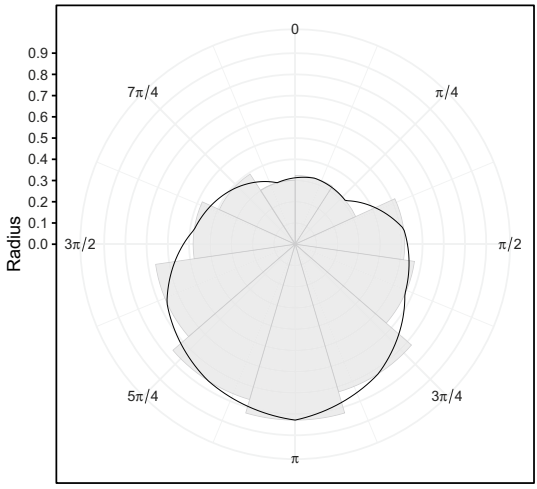


Fig. 5 Polygon rose diagram is the linear interpolation of the mid-points of a sector in the rose diagram

Theorem 3 *Assume that* (A1), (A2), (A4), *and* (A5) *hold. Then*,

$$\mathrm{MISE}[\tilde{r}_f(\cdot;h)] = \mathrm{AMISE}[\tilde{r}_f(\cdot;h)] + o(h^4 + (nh)^{-1}),$$

where $\mathrm{AMISE}[\tilde{r}(\theta;h)]$ *is*:

$$\mathrm{AMISE}[\tilde{r}_f(\cdot;h)] = G(r_f'', r_f', r_f)h^4 + \frac{2\pi}{3nh}, \tag{14}$$

where $G(r_f'', r_f', r_f) := 49R(r_f'')/2880 + R((r_f'')^{1/2} r_f' r_f^{-1/2})/96 + R((r_f')^2 r_f^{-1})/576$. *The optimizer of the RHS in Eq.* (14) *is*:

$$h_{\mathrm{poly}} = [\pi/6G(r_f'', r_f', r_f)]^{1/5} n^{-1/5}. \tag{15}$$

Combining (14) *and* (15) *provide* $\mathrm{MISE}[\tilde{r}_f(\cdot;h_{\mathrm{poly}})] = O(n^{-4/5})$.

We present the proof in Appendix E. The convergence rate in the MISE is equivalent to that of the circular kernel density estimator and the frequency polygon. Moreover, the polygon rose diagram estimator outperforms the rose diagram estimator.

We consider a new bin width estimator for this estimator by employing the von Mises density as the reference density. The von Mises density has:

$$G(r_f'', r_f', r_f) = [47\kappa^2 I_2(\kappa)/I_0(\kappa)] + 49\kappa^2]/11520. \tag{16}$$

We then propose the following NRR for the polygon rose diagram estimator:

Definition 4

$$\hat{h}_{\mathrm{PNR}} = \{1920\pi[47\hat{\kappa}^2 I_2(\hat{\kappa})/I_0(\hat{\kappa}) + 49\hat{\kappa}^2]^{-1}\}^{1/5} n^{-1/5},$$

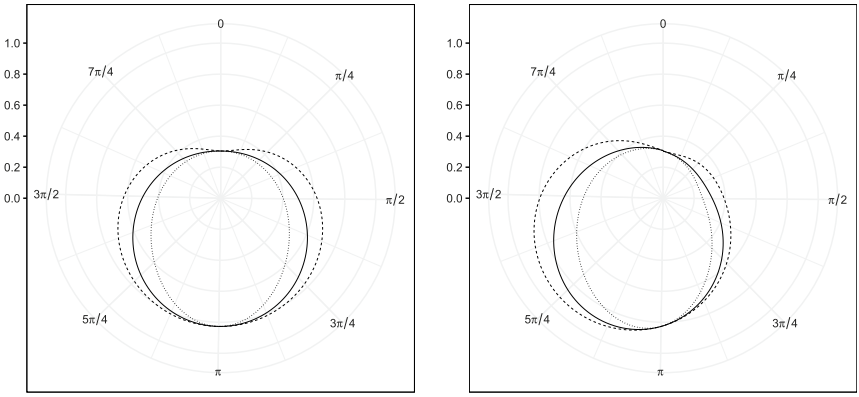
where $\hat{\kappa}$ is the MLE of κ.

From Algorithm 1, we present the ENRR for this by employing the AMISE that estimates $G(r_f'', r_f', r_f)$ in the RHS of Eq. (14) by Eq. (16).

4 Numerical experiment

4.1 Comparison between bin width selection rules for rose diagram

We now investigate how a choice of the bin width affects the performance of the rose diagram under finite samples. We perform a simulation to compare the proposed bin width estimator in Sect. 2 to the rules from previous literature stated in Sect. 1 under finite sample sizes. Sections 1 and 2 put the starting point at zero without losing generality. We also want to know how the selection of the starting point may affect its performance with small samples. As the starting point, we choose the sample circular mean $\bar{\theta}$, the opposite point $\bar{\theta}_\pi := \bar{\theta} + \pi \pmod{2\pi}$, the right point



(a) Models 1–3. (b) Models 4–6.

Fig. 6 Each line denotes the square root of twice the density. **a** Models 1–3 are von Mises (solid line), cardioid (dashed line), and wrapped Cauchy (dotted line) distributions. **b** Models 4–6 are sine-skewed von Mises(solid line), sine-skewed cardioid (dashed line), and sine-skewed wrapped Cauchy distributions(dotted line)

$\bar{\theta}_{\mathrm{R}} := \bar{\theta} + \pi/2 \pmod{2\pi}$, and the left point $\bar{\theta}_{\mathrm{L}} := \bar{\theta} + 3\pi/2 \pmod{2\pi}$. For example, in the data generated from symmetric unimodal distributions, the mean $\bar{\theta}$ is the concentrated position and the opposite $\bar{\theta}_{\pi}$ is the sparse position.

We conduct the numerical experiment under the six simulation scenarios Models 1–6, as presented in Fig. 6, which were employed by Tsuruta and Sagae (2020). Model 1 (the von Mises distribution), Model 2 (the cardioid distribution), and Model 3 (the wrapped Cauchy distribution) are well-used in numerical experiments, whereas Model 4 (the sine-skewed von Mises distribution), Model 5 (the sine-skewed cardioid distribution), and Model 6 (the sine-skewed wrapped Cauchy distribution) are produced by skewing the first three distributions. These six distributions are a subclass of a sine-skewed Jones–Pewsey distribution (Abe and Pewsey 2011).

We define the previous literature's rules as the following:

(R1) Ballantyne and Cornish's rule (BCR) is given by: $\hat{h}_{\mathrm{BCR}} := 2\pi/g_{\mathrm{BCR}}(n)$ with $g_{\mathrm{BCR}}(n) := n/5$ (Ballantyne and Cornish 1979).

(R2) Andreasen's rule (AR) is given by:

$$\hat{h}_{\mathrm{AR}} := 2 \times \mathrm{tdf} \times \frac{\mathrm{CSD}}{n^{1/2}},$$

where tdf denotes Student's t-value for $n-1$ degrees of freedom at the.95 level and CSD is circular standard deviation (Andreasen 1990).

(R3) Square root rule (SRR) is given by: $\hat{h}_{\mathrm{SRR}} := 2\pi/g_{\mathrm{SRR}}(n)$ with $g_{\mathrm{SRR}}(n) := n^{1/2}$ (Pewsey et al. 2013).

(R4) Sanderson and Peacock's rule (SPR) is given by: $\hat{h}_{\mathrm{SPR}} := 2\pi/g_{\mathrm{SPR}}(n)$ with $g_{\mathrm{SPR}}(n) := 2n^{1/3}$(Sanderson and Peacock 2020).

For Rules (R2)–(R4), we consider the need to transform these rules to all equal bin widths as Sanderson and Peacock (2020) rounded up $g_{\mathrm{SPR}}(n)$ to the nearest integer, although BCR provides all equal bin widths because all sample sizes are multiples of five. Therefore, we also employ the following rules:

(R5) Equal Andreasen's rule (EAR) is given by: $\hat{h}_{\mathrm{EAR}} := 2\pi/\lceil g_{\mathrm{AR}}(n) \rceil$ with $g_{\mathrm{AR}}(n) := 2\pi/\hat{h}_{\mathrm{AR}}$.
(R6) Equal square root rule (ESRR) is given by: $\hat{h}_{\mathrm{ESRR}} := 2\pi/\lceil g_{\mathrm{SRR}}(n) \rceil$.
(R7) Equal Sanderson and Peacock's rule (ESPR) is given by: $\hat{h}_{\mathrm{ESPR}} := 2\pi/\lceil g_{\mathrm{SPR}}(n) \rceil$.

Note that AR and EAR are only used in unimodal distributions. We execute the simulation with `R` according to the following procedure:

1. Execute the following steps for Model 1:
   (a) Generate a random sample of size n distributed in Model 1.
   (b) Calculate each rose diagram's numerical ISE: $\mathrm{ISE}_{\mathrm{rose}} := \int_0^{2\pi} \{\hat{r}_f(\theta;\hat{h}) - r_f(\theta)\}^2 d\theta$, with any bin width estimator as $\hat{h}$.
   (c) Repeat (a)–(b) 1,000 times and then, calculate each mean and standard error of $\mathrm{ISE}_{\mathrm{rose}}$
2. Execute steps (a)–(c) for Models 2–6.

We now discuss the results of the numerical experiment. Tables 1 and 2 show little difference between all the methods for selecting the starting point in all the bin width selectors in Models 1–6 when $n \geq 500$ (Supplementary Table 6 also shows this future in all the bin width selectors in all the scenarios when $n = 200$). The effect of choosing the starting point is significantly less than that of the choice of bin width in all the models in all sample sizes.

The MISE of BCR is hardly decreasing in all cases. This is because the order of BCR is n^{-1}, but the order of n^{-1} does not have consistency in Theorem 1. When $n \geq 500$, the selectors of order $n^{-1/2}$ (AR, SRR, EAR, and ESPR) perform less than the selectors of order $n^{-1/3}$ (NRR, ENRR, BCV, EBCV, SPR, and ESPR), in all the models. This indicates that the optimal magnitude of the bin widths is $n^{-1/3}$ as (6). BCV and EBCV outperform SPR and ESPR in all models when $n \geq 500$. When $n \geq 500$, although NRR, ENRR, SPR, and ESPR have little difference in Models 3 and 6, NRR and ENRR outperform SPR and EPSPR in the other scenarios. This shows that we should appropriately estimate the quantity of the roughness $R(r_f')$ in Eq. (6).

BCV has the best performance, and EBCV has the second-best performance in all the scenarios when $n \geq 500$. Further, when $n \geq 500$, the third performances are NRR and ENRR, although NRR and ENRR have little difference in performance from SPR and ESPR in Models 3 and 6.

Now, we discuss the effect of the starting point in small samples. Supplementary Tables 4 and 5 show that the right point of EBCV outperforms the other

Table 1 Means (its standard error) of 100×integrated squared error of the rose diagrams based on the number of repetitions N=1000 in Models 1–6 in the simulation in Sect. 4. The sample size is $n = 500$. In the starting point t_0, M denotes the sample circular mean, O denotes the opposite, R denotes the right point, and L denotes the left point

Selector	t_0	Model1	Model2	Model3	Model4	Model5	Model6
NRR	M	1.68(.02)	1.54(.01)	1.92(.01)	1.95(.02)	1.79(.02)	2.16(.02)
NRR	O	1.71(.02)	1.54(.02)	1.97(.02)	1.96(.02)	1.80(.02)	2.18(.02)
NRR	R	1.67(.02)	1.51(.01)	1.94(.02)	1.92(.02)	1.75(.02)	2.18(.02)
NRR	L	1.67(.02)	1.53(.02)	1.94(.02)	1.93(.02)	1.77(.02)	2.15(.02)
ENRR	M	1.64(.02)	1.50(.01)	1.91(.01)	1.89(.02)	1.72(.02)	2.12(.02)
ENRR	O	1.64(.02)	1.49(.01)	1.90(.01)	1.89(.02)	1.73(.02)	2.11(.02)
ENRR	R	1.63(.02)	1.47(.01)	1.89(.01)	1.88(.02)	1.72(.02)	2.12(.02)
ENRR	L	1.64(.02)	1.48(.01)	1.90(.01)	1.89(.02)	1.72(.02)	2.10(.02)
BCV	M	1.40(.01)	1.30(.01)	1.64(.01)	1.61(.01)	1.48(.01)	1.83(.01)
BCV	O	1.41(.01)	1.33(.02)	1.66(.01)	1.62(.02)	1.47(.01)	1.86(.02)
BCV	R	1.37(.01)	1.29(.02)	1.68(.01)	1.56(.01)	1.41(.01)	1.83(.02)
BCV	L	1.37(.01)	1.26(.01)	1.79(.04)	1.54(.01)	1.45(.01)	1.82(.02)
EBCV	M	1.51(.01)	1.41(.01)	1.77(.01)	1.74(.01)	1.62(.01)	1.95(.01)
EBCV	O	1.53(.01)	1.46(.03)	1.80(.02)	1.80(.03)	1.61(.01)	1.99(.02)
EBCV	R	1.54(.02)	1.44(.02)	1.80(.01)	1.73(.01)	1.60(.01)	1.96(.02)
EBCV	L	1.53(.01)	1.41(.01)	1.93(.03)	1.73(.01)	1.61(.01)	1.99(.02)
BCR	M	13.35(.07)	13.03(.07)	13.54(.07)	13.83(.07)	13.76(.07)	13.89(.07)
BCR	O	13.35(.07)	13.03(.07)	13.54(.07)	13.83(.07)	13.76(.07)	13.89(.07)
BCR	R	13.35(.07)	13.03(.07)	13.54(.07)	13.83(.07)	13.76(.07)	13.89(.07)
BCR	L	13.35(.07)	13.03(.07)	13.54(.07)	13.83(.07)	13.76(.07)	13.89(.07)
AR	M	3.03(.03)	2.73(.03)	3.14(.03)	3.55(.03)	3.30(.03)	3.54(.03)
AR	O	3.08(.03)	2.80(.03)	3.13(.03)	3.58(.03)	3.31(.03)	3.57(.03)
AR	R	3.04(.03)	2.76(.03)	3.11(.03)	3.56(.03)	3.27(.03)	3.56(.03)
AR	L	3.03(.03)	2.78(.03)	3.15(.03)	3.52(.03)	3.30(.03)	3.55(.03)
SRR	M	2.44(.02)	2.42(.02)	2.50(.02)	2.55(.02)	2.52(.02)	2.60(.02)
SRR	O	2.53(.02)	2.49(.02)	2.52(.02)	2.61(.03)	2.54(.02)	2.67(.03)
SRR	R	2.42(.02)	2.38(.02)	2.53(.02)	2.58(.03)	2.48(.02)	2.66(.03)
SRR	L	2.42(.02)	2.41(.02)	2.51(.02)	2.59(.03)	2.5(.02)	2.66(.03)
SPR	M	1.85(.02)	1.78(.02)	1.95(.02)	2.01(.02)	1.88(.02)	2.15(.02)
SPR	O	1.84(.02)	1.77(.02)	1.97(.02)	2.00(.02)	1.88(.02)	2.14(.02)
SPR	R	1.83(.02)	1.78(.02)	1.96(.02)	2.00(.02)	1.86(.02)	2.14(.02)
SPR	L	1.83(.02)	1.77(.02)	1.96(.02)	1.98(.02)	1.88(.02)	2.14(.02)
EAR	M	3.05(.03)	2.75(.03)	3.13(.03)	3.53(.03)	3.31(.03)	3.52(.03)
EAR	O	3.05(.03)	2.75(.03)	3.12(.03)	3.56(.03)	3.30(.03)	3.53(.03)
EAR	R	3.06(.03)	2.77(.03)	3.12(.03)	3.55(.03)	3.30(.03)	3.53(.03)
EAR	L	3.08(.03)	2.78(.03)	3.12(.03)	3.55(.03)	3.27(.03)	3.54(.03)
ESRR	M	2.42(.02)	2.38(.02)	2.51(.02)	2.58(.03)	2.50(.02)	2.63(.03)
ESRR	O	2.42(.02)	2.42(.02)	2.49(.02)	2.55(.02)	2.47(.02)	2.59(.02)
ESRR	R	2.45(.02)	2.40(.02)	2.46(.02)	2.56(.03)	2.45(.02)	2.61(.02)
ESRR	L	2.44(.02)	2.39(.02)	2.50(.02)	2.55(.02)	2.52(.02)	2.60(.02)

Table 1 (continued)

Selector	t_0	Model1	Model2	Model3	Model4	Model5	Model6
ESPR	M	1.84(.02)	1.78(.02)	1.97(.02)	2.00(.02)	1.88(.02)	2.14(.02)
ESPR	O	1.84(.02)	1.78(.02)	1.97(.02)	2.00(.02)	1.88(.02)	2.14(.02)
ESPR	R	1.84(.02)	1.78(.02)	1.97(.02)	2.00(.02)	1.88(.02)	2.14(.02)
ESPR	L	1.84(.02)	1.78(.02)	1.97(.02)	2.00(.02)	1.88(.02)	2.14(.02)

points in Models 1, 3, 4, and 5 when $n \leq 100$. The right point is placed in the direction of the skewness of the underlying distribution in Models 4 and 5. The right point of BCV has better performance than the other points in Models 1 and 5 when $n \leq 100$. NRR and ENRR have a smaller effect on the starting points when $n \leq 100$, as NRR and ENRR do not depend on the starting point.

4.2 Comparison rose diagram and polygon rose diagram estimators

We now compare the rose diagram and polygon rose diagram estimators in Models 1–6 under finite sample sizes. We also use NRR, ENRR, BCV, and EBCV as the selectors for the rose diagram. The selectors for the polygon rose diagram are NRR and ENRR as defined in Sect. 3. The same methods are used for selecting the starting point as in Sect. 4.1.

We now discuss the result of the numerical experiment. Table 3 shows that the polygon rose diagram estimators (with NRR or ENRR) outperform the rose diagram estimators (with NRR, ENRR, BCV, or EBCV) in all models when $n \geq 500$. However, the result could depend on the bin width selection. ENRR performs better than NRR in all models in the polygon rose diagram estimators when $n = 500$. Moreover, when $n = 1000$, ENRR outperforms NRR in Models 1,2, 3, and 5.

We also summarize the effect of the starting point on the polygon rose diagram. Supplementary Tables 4, 5, 6, 7, and 8 show that ENRR has a smaller effect on the starting points than NRR. ENRR is largely unaffected by the starting point when $n \geq 100$ and outperforms NRR in almost all models when $n \leq 500$. Therefore, we recommend using ENRR for the polygon rose diagram.

5 Application

We now illustrate a real-data example showing a rose diagram with the proposed bin width selectors (ENRR and EBCV). We choose these selectors because software often requires inputting the number of the bins under all equal bin widths such as `rose.diag` in the `circular` library (Agostinelli and Lund 2017) in statistical software `R` (R Core Team 2021). We use `wind` data in the `circular` library (see Fig. 7). In `wind` data, the sample circular mean $\bar{\theta}$ is 0.2922, the mean resultant length is 0.6557, and the skewness is 0.3493. For the starting point, we employ the

Table 2 Means (its standard error) of 100×integrated squared error of the rose diagrams based on the number of repetitions N=1000 in Models 1–6 in the simulation in Sect. 4. The sample size is $n = 1000$. In the starting point t_0, M denotes the sample circular mean, O denotes the opposite, R denotes the right point, and L denotes the left point

Selector	t_0	Model1	Model2	Model3	Model4	Model5	Model6
NRR	M	1.04(.01)	.97(.01)	1.22(.01)	1.20(.01)	1.14(.01)	1.35(.01)
NRR	O	1.04(.01)	.98(.01)	1.24(.01)	1.22(.01)	1.15(.01)	1.36(.01)
NRR	R	1.03(.01)	.95(.01)	1.22(.01)	1.19(.01)	1.13(.01)	1.35(.01)
NRR	L	1.03(.01)	.96(.01)	1.23(.01)	1.20(.01)	1.12(.01)	1.35(.01)
ENRR	M	1.02(.01)	.94(.01)	1.20(.01)	1.18(.01)	1.11(.01)	1.33(.01)
ENRR	O	1.02(.01)	.94(.01)	1.21(.01)	1.17(.01)	1.11(.01)	1.32(.01)
ENRR	R	1.01(.01)	.94(.01)	1.20(.01)	1.17(.01)	1.12(.01)	1.32(.01)
ENRR	L	1.02(.01)	.94(.01)	1.20(.01)	1.18(.01)	1.10(.01)	1.33(.01)
BCV	M	.86(.01)	.79(.01)	1.05(.01)	1.03(.01)	.95(.01)	1.17(.01)
BCV	O	.86(.01)	.80(.01)	1.06(.01)	1.03(.01)	.96(.01)	1.18(.01)
BCV	R	.84(.01)	.77(.01)	1.06(.01)	1.02(.01)	.92(.01)	1.19(.01)
BCV	L	.85(.01)	.78(.01)	1.06(.01)	1.01(.01)	.94(.01)	1.17(.01)
EBCV	M	.94(.01)	.87(.01)	1.12(.01)	1.12(.01)	1.03(.01)	1.26(.01)
EBCV	O	.94(.01)	.87(.01)	1.14(.01)	1.12(.01)	1.04(.01)	1.27(.01)
EBCV	R	.93(.01)	.86(.01)	1.14(.01)	1.11(.01)	1.04(.01)	1.27(.01)
EBCV	L	.94(.01)	.87(.01)	1.14(.01)	1.11(.01)	1.04(.01)	1.27(.01)
BCR	M	13.33(.05)	12.98(.05)	13.47(.05)	13.94(.05)	13.7(.05)	13.87(.05)
BCR	O	13.33(.05)	12.98(.05)	13.47(.05)	13.94(.05)	13.7(.05)	13.87(.05)
BCR	R	13.33(.05)	12.98(.05)	13.47(.05)	13.94(.05)	13.7(.05)	13.87(.05)
BCR	L	13.33(.05)	12.98(.05)	13.47(.05)	13.94(.05)	13.7(.05)	13.87(.05)
AR	M	2.07(.02)	1.89(.01)	2.15(.02)	2.37(.02)	2.23(.02)	2.33(.02)
AR	O	2.08(.02)	1.90(.02)	2.14(.02)	2.40(.02)	2.24(.02)	2.35(.02)
AR	R	2.08(.02)	1.87(.01)	2.15(.02)	2.38(.02)	2.24(.02)	2.34(.02)
AR	L	2.07(.02)	1.87(.01)	2.15(.01)	2.37(.02)	2.24(.02)	2.35(.02)
SRR	M	1.65(.01)	1.62(.01)	1.69(.01)	1.71(.01)	1.69(.01)	1.74(.01)
SRR	O	1.65(.01)	1.64(.01)	1.69(.01)	1.73(.01)	1.70(.01)	1.72(.01)
SRR	R	1.65(.01)	1.63(.01)	1.69(.01)	1.71(.01)	1.69(.01)	1.72(.01)
SRR	L	1.64(.01)	1.62(.01)	1.69(.01)	1.72(.01)	1.71(.01)	1.73(.01)
SPR	M	1.13(.01)	1.10(.01)	1.24(.01)	1.23(.01)	1.19(.01)	1.32(.01)
SPR	O	1.13(.01)	1.10(.01)	1.24(.01)	1.23(.01)	1.19(.01)	1.32(.01)
SPR	R	1.13(.01)	1.10(.01)	1.24(.01)	1.23(.01)	1.19(.01)	1.32(.01)
SPR	L	1.13(.01)	1.10(.01)	1.24(.01)	1.23(.01)	1.19(.01)	1.32(.01)
EAR	M	2.07(.02)	1.88(.01)	2.15(.02)	2.38(.02)	2.24(.02)	2.34(.02)
EAR	O	2.07(.02)	1.89(.02)	2.15(.02)	2.39(.02)	2.23(.02)	2.35(.02)
EAR	R	2.06(.02)	1.89(.01)	2.14(.02)	2.37(.02)	2.23(.02)	2.36(.02)
EAR	L	2.07(.02)	1.88(.01)	2.13(.02)	2.37(.02)	2.24(.02)	2.36(.02)
ESRR	M	1.64(.01)	1.63(.01)	1.69(.01)	1.72(.01)	1.69(.01)	1.72(.01)
ESRR	O	1.64(.01)	1.63(.01)	1.69(.01)	1.72(.01)	1.69(.01)	1.72(.01)
ESRR	R	1.64(.01)	1.63(.01)	1.69(.01)	1.72(.01)	1.69(.01)	1.72(.01)
ESRR	L	1.64(.01)	1.63(.01)	1.69(.01)	1.72(.01)	1.69(.01)	1.72(.01)

Table 2 (continued)

Selector	t_0	Model1	Model2	Model3	Model4	Model5	Model6
ESPR	M	1.13(.01)	1.10(.01)	1.24(.01)	1.23(.01)	1.19(.01)	1.32(.01)
ESPR	O	1.13(.01)	1.10(.01)	1.24(.01)	1.23(.01)	1.19(.01)	1.32(.01)
ESPR	R	1.13(.01)	1.10(.01)	1.24(.01)	1.23(.01)	1.19(.01)	1.32(.01)
ESPR	L	1.13(.01)	1.10(.01)	1.24(.01)	1.23(.01)	1.19(.01)	1.32(.01)

right point according to the direction of the skewness. Figure 8 shows that two rose diagrams with ENRR or EBCV represent that the distribution has the mode between north and northeast and is almost symmetric, although the jumps of the rose diagrams obscure the existence of small modes. Figure 8 indicates that the polygon rose diagram with ENRR has the mode between north and northeast and is smoother than the rose diagrams.

6 Conclusion

In this study, we pointed out that the radius of the rose diagram (the square root of the twice density) is a nonparametric estimator of the square root of twice density. We set a condition that the $m-1$ bin widths is a common bin width h, and m-th bin width is ch with $c \in (0, 1]$. Then, we employed MISE as the global error criterion between the rose diagram and the square root of twice density. We showed that choosing the optimal bin width of order $n^{-1/3}$ provides the minimizing MISE of order $n^{-2/3}$. Moreover, we provided an asymptotic normal distribution of its estimator with equal variance when $c = 1$; that is, this means employing all the equal bin widths, although the histogram has a heterogeneous variance that depends on an underlying density. This property is an advantage of the rose diagram estimator.

Next, we proposed some data-based bin width selection rules. NRR employs the von Mises density as the underlying density, while BCV is the nonparametric estimator that does not assume a specific density. These rules often show that one bin width is shorter than the other bin width. In this practice, we assumed that statisticians prefer all equal bin widths to avoid arbitrariness in the placement of the shorter bin width. Further, software of rose diagrams often requires all equal bin widths. We also provided ENRR and EBCV as equal split bin width estimators. The numerical experiment showed that BCV and EBCV outperform NRR and ENRR under the finite samples, although our four rules have better performance than all the rules in the literature introduced in Sect. 1. Moreover, the effect of choosing the starting point is significantly less than that of the choice of bin width in all scenarios. In particular, when sample size $n \geq 200$, we can ignore the effect of the starting point on the rose diagram.

We believed that all equal split bin width selectors such as ENRR and EBCV have the following three advantages: First, these selectors avoid placing one bin width with different lengths than the others. Second, all the equal bin widths provide

Table 3 Means (its standard error) of 100×integrated squared error of the rose diagrams (denoted by Rose) with NRR, ENRR, BCV, or EBCV as its bin width selector) and the polygon rose diagrams (denoted by Poly) with NRR or ENRR based on the number of repetitions N=1000 in Models 1–6 in the simulation in Sect. 4. The sample size is $n = 500$. In the starting point t_0, M denotes the sample circular mean, O denotes the opposite, R denotes the right point, and L denotes the left point. The top block represents MISEs in sample size $n = 500$ and the bottom block represents MISEs in sample size $n = 1000$

$\hat{r}_f$	Selector	t_0	Model1	Model2	Model3	Model4	Model5	Model6
Rose	NRR	M	1.68(.02)	1.54(.01)	1.92(.01)	1.95(.02)	1.79(.02)	2.16(.02)
Rose	NRR	O	1.71(.02)	1.54(.02)	1.97(.02)	1.96(.02)	1.80(.02)	2.18(.02)
Rose	NRR	R	1.67(.02)	1.51(.01)	1.94(.02)	1.92(.02)	1.75(.02)	2.18(.02)
Rose	NRR	L	1.67(.02)	1.53(.02)	1.94(.02)	1.93(.02)	1.77(.02)	2.15(.02)
Rose	ENRR	M	1.64(.02)	1.5(.01)	1.91(.01)	1.89(.02)	1.72(.02)	2.12(.02)
Rose	ENRR	O	1.64(.02)	1.49(.01)	1.90(.01)	1.89(.02)	1.73(.02)	2.11(.02)
Rose	ENRR	R	1.63(.02)	1.47(.01)	1.89(.01)	1.88(.02)	1.72(.02)	2.12(.02)
Rose	ENRR	L	1.64(.02)	1.48(.01)	1.90(.01)	1.89(.02)	1.72(.02)	2.10(.02)
Rose	BCV	M	1.40(.01)	1.30(.01)	1.64(.01)	1.61(.01)	1.48(.01)	1.83(.01)
Rose	BCV	O	1.41(.01)	1.33(.02)	1.66(.01)	1.62(.02)	1.47(.01)	1.86(.02)
Rose	BCV	R	1.37(.01)	1.29(.02)	1.68(.01)	1.56(.01)	1.41(.01)	1.83(.02)
Rose	BCV	L	1.37(.01)	1.26(.01)	1.79(.04)	1.54(.01)	1.45(.01)	1.82(.02)
Rose	EBCV	M	1.51(.01)	1.41(.01)	1.77(.01)	1.74(.01)	1.62(.01)	1.95(.01)
Rose	EBCV	O	1.53(.01)	1.46(.03)	1.80(.02)	1.80(.03)	1.61(.01)	1.99(.02)
Rose	EBCV	R	1.54(.02)	1.44(.02)	1.80(.01)	1.73(.01)	1.60(.01)	1.96(.02)
Rose	EBCV	L	1.53(.01)	1.41(.01)	1.93(.03)	1.73(.01)	1.61(.01)	1.99(.02)
Poly	NRR	M	.71(.03)	.89(.06)	.93(.06)	.95(.07)	.95(.06)	1.05(.06)
Poly	NRR	O	.75(.02)	.75(.02)	1.02(.05)	.99(.06)	.89(.02)	1.19(.02)
Poly	NRR	R	.74(.02)	1.17(.36)	.96(.03)	.99(.05)	1.00(.06)	1.18(.04)
Poly	NRR	L	.75(.03)	.84(.04)	1.03(.03)	.94(.03)	.94(.04)	1.08(.03)
Poly	ENRR	M	.61(.01)	.56(.01)	.87(.01)	.69(.01)	.62(.01)	.87(.01)
Poly	ENRR	O	.61(.01)	.56(.01)	.83(.01)	.70(.01)	.62(.01)	.95(.01)
Poly	ENRR	R	.61(.01)	.56(.01)	.87(.01)	.69(.01)	.62(.01)	.93(.01)
Poly	ENRR	L	.61(.01)	.57(.01)	.83(.01)	.69(.01)	.62(.01)	.91(.01)
Rose	NRR	M	1.04(.01)	.97(.01)	1.22(.01)	1.20(.01)	1.14(.01)	1.35(.01)
Rose	NRR	O	1.04(.01)	.98(.01)	1.24(.01)	1.22(.01)	1.15(.01)	1.36(.01)
Rose	NRR	R	1.03(.01)	.95(.01)	1.22(.01)	1.19(.01)	1.13(.01)	1.35(.01)
Rose	NRR	L	1.03(.01)	.96(.01)	1.23(.01)	1.20(.01)	1.12(.01)	1.35(.01)
Rose	ENRR	M	1.02(.01)	.94(.01)	1.20(.01)	1.18(.01)	1.11(.01)	1.33(.01)
Rose	ENRR	O	1.02(.01)	.94(.01)	1.21(.01)	1.17(.01)	1.11(.01)	1.32(.01)
Rose	ENRR	R	1.01(.01)	.94(.01)	1.20(.01)	1.17(.01)	1.12(.01)	1.32(.01)
Rose	ENRR	L	1.02(.01)	.94(.01)	1.20(.01)	1.18(.01)	1.10(.01)	1.33(.01)
Rose	BCV	M	.86(.01)	.79(.01)	1.05(.01)	1.03(.01)	.95(.01)	1.17(.01)
Rose	BCV	O	.86(.01)	.80(.01)	1.06(.01)	1.03(.01)	.96(.01)	1.18(.01)
Rose	BCV	R	.84(.01)	.77(.01)	1.06(.01)	1.02(.01)	.92(.01)	1.19(.01)
Rose	BCV	L	.85(.01)	.78(.01)	1.06(.01)	1.01(.01)	.94(.01)	1.17(.01)
Rose	EBCV	M	.94(.01)	.87(.01)	1.12(.01)	1.12(.01)	1.03(.01)	1.26(.01)
Rose	EBCV	O	.94(.01)	.87(.01)	1.14(.01)	1.12(.01)	1.04(.01)	1.27(.01)

Table 3 (continued)

$\hat{r}_f$	Selector	t_0	Model1	Model2	Model3	Model4	Model5	Model6
Rose	EBCV	R	.93(.01)	.86(.01)	1.14(.01)	1.11(.01)	1.04(.01)	1.27(.01)
Rose	EBCV	L	.94(.01)	.87(.01)	1.14(.01)	1.11(.01)	1.04(.01)	1.27(.01)
Poly	NRR	M	.40(.02)	.80(.21)	.62(.04)	.47(.03)	.65(.05)	.53(.02)
Poly	NRR	O	.42(.01)	.48(.02)	.70(.03)	.46(.01)	.56(.02)	.59(.01)
Poly	NRR	R	.42(.02)	.54(.03)	.67(.03)	.46(.01)	.56(.02)	.63(.02)
Poly	NRR	L	.43(.02)	.55(.04)	.67(.03)	.45(.01)	.61(.03)	.59(.01)
Poly	ENRR	M	.35(.01)	.32(.01)	.52(.01)	.41(.01)	.38(.01)	.55(.01)
Poly	ENRR	O	.35(.01)	.33(.01)	.52(.01)	.41(.01)	.38(.01)	.54(.01)
Poly	ENRR	R	.34(.01)	.33(.01)	.49(.01)	.40(.01)	.37(.01)	.54(.01)
Poly	ENRR	L	.34(.01)	.32(.01)	.49(.01)	.40(.01)	.37(.01)	.56(.01)

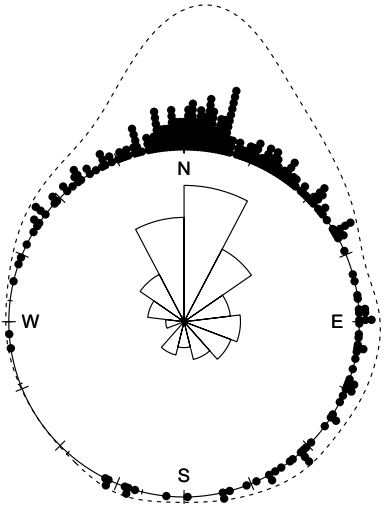
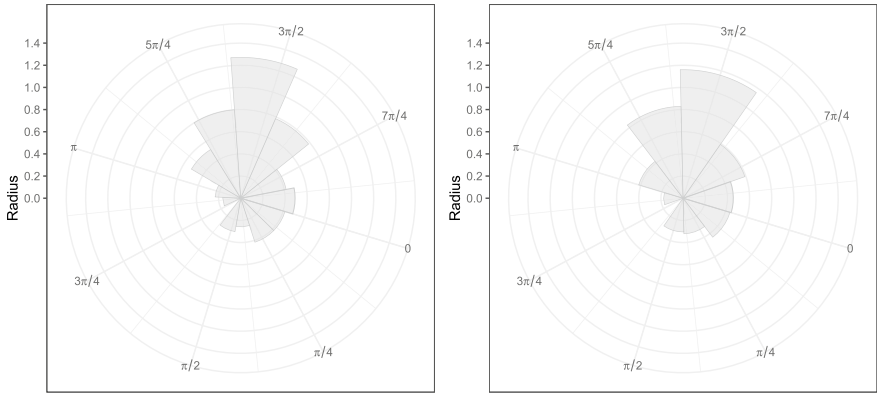


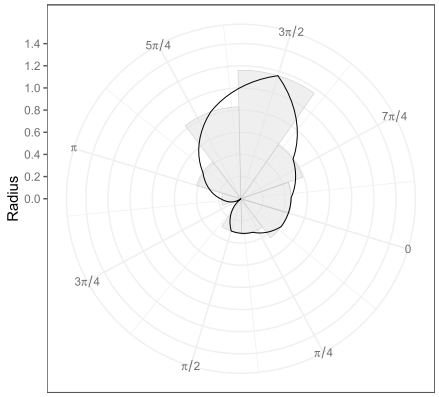
Fig. 7 Circular wind data plot of `wind` data ($n = 310$), which were recorded at a meteorological station named "Col de la Roa" in the Italian Alps from January 29, 2001 to March 31, 2001. The values are measured in radians, clockwise from north. The rose diagram is illustrated by `rose.diag` with bin width 0.4833 estimated by ENRR (the number of the bins is thirteen). The dashed line is circular kernel density estimation with bandwidth 32.51 estimated by the direct plug-in selector proposed by Tsuruta and Sagae (2020)

homoskedasticity. Third, we can employ these selectors in commonly used software such as `R` because software often requires inputting the number of the bins under all equal bin widths. Therefore, we recommend employing EBCV because it has the best performance in the equal split bin width selectors in the numerical experiment. When employing EBCV, if $n \geq 200$, we observed little difference in the choice of the starting points in the numerical experiment. Whereas, if $n \leq 100$, then, we should select either the right point or the left point as the starting point because these points have often better performance than the other points in the numerical experiment. We also recommend that the choice between the two points should be made according to the direction of skewness of the underlying distribution.

We presented a polygon rose diagram estimator to improve the rose diagram with linear interpolation. We obtained the convergence rate of the optimal MISE of the polygon rose diagram as $O(n^{-4/5})$. The numerical experiment showed that



(a) Rose diagram (ENRR) (b) Rose diagram (EBCV)



(c) Polygon rose diagram (ENRR)

Fig. 8 **a** The rose diagram with bin width 0.4833 estimated by ENRR. **b** The rose diagram with bin width 0.6283 estimated by EBCV. **c** The solid line represents the polygon rose diagram with bin width 0.6283 estimated by ENRR. Each starting point in (**a**)–(**c**) is denoted by zero. The values in the circle are measured clockwise from the starting point. All the starting points are 1.863 (the right point), which is the radian measured from the north (upper direction)

the polygon rose diagram performs better than the rose diagram in all the scenarios, although this result may depend on the bin width choice.

For further studies, we will derive the convergence rate of the BCV estimator. We believe that the result can provide a comparison of its performance with that of any other estimator. The problem of the starting point can be solved by the smoothness, such as a moving average rose diagram, as proposed by Munro and Blenkinsop (2012). Therefore, as a future objective, we shall investigate the theoretical property of a moving average rose diagram with smoothness.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10463-023-00868-4.

Acknowledgements We would like to thank the reviewers for their helpful comments. This work was supported by JSPS KAKENHI Grant Numbers JP20K19760 and JP21K11782.

Appendix A

Proof of Theorem 1 Recalling that $W_k = \hat{f}(\theta;h_k) - \mathrm{E}[\hat{f}(\theta;h_k)]$ for θ in B_k, the Taylor expansion of the rose diagram estimator $r_f(\theta;h)$ around $\mathrm{E}[\hat{f}(\theta;h_k)]$ is:

$$\hat{r}_f(\theta;h) = \sqrt{2\mathrm{E}[\hat{f}(\theta;h_k)]} + \sum_{t=1}^{\infty}(-1)^{t-1}\frac{\sqrt{2}(2t-3)!!}{2^t t!}\mathrm{E}[\hat{f}(\theta;h_k)]^{1/2-t}W_k^t, \tag{17}$$

for $\theta \in B_k$, where $(2t-1)!! := 1\cdot 3\cdot 5\cdots(2t-3)(2t-1)$ and (-1)!!:=1. Combining Lemma 1 and (17), the expectation of the rose diagram estimator approximates to:

$$\mathrm{E}[\hat{r}_f(\theta;h)] = \sqrt{2p_k/h_k} + O((nh)^{-1}),\ \theta \in B_k. \tag{18}$$

Note that $B_k = [t_k, t_k + h_k)$. From the Taylor expansion of $f(\cdot)$ around $\theta \in B_k$, we obtain

$$\begin{aligned} p_k &= \int_{t_k}^{t_k+h_k} f(y)dy \\ &= \int_{t_k}^{t_k+h_k} [f(\theta) + f'(\theta)(y-\theta) + O((y-\theta)^2)]dy \\ &= (t_k+h_k)f(\theta) + f'(\theta)(t_k+h_k-\theta)^2/2 + O((t_k+h_k-\theta)^3) \\ &\quad - [t_k f(\theta) + f'(\theta)(t_k-\theta)^2/2 + O((t_k-\theta)^3)] \\ &= h_k f(\theta) + f'(\theta)[h_k^2/2 + (t_k-\theta)h_k] + O(h^3). \end{aligned} \tag{19}$$

Approximating p_k in the RHS in Eq. (18) by Eq. (19) leads to:

$$\begin{aligned} \mathrm{E}[\hat{r}_f(\theta;h)] &= \sqrt{2}\sqrt{f(\theta) + f'(\theta)(h_k/2 + t_k - \theta) + O(h^2)} + O((nh^{-1})) \\ &= \sqrt{2f(\theta)} + (2f(\theta))^{-1/2}f'(\theta)(h_k/2 + t_k - \theta) + O(h^2 + (nh)^{-1}) \\ &= r_f(\theta) + r_f'(\theta)(h_k/2 + t_k - \theta) + O(h^2 + (nh)^{-1}), \end{aligned} \tag{20}$$

for $\theta \in B_k$. From (20), we obtain:

$$\mathrm{Bias}[\hat{r}_f(\theta;h)] = r_f'(\theta)(h_k/2 + t_k - \theta) + O(h^2 + (nh)^{-1}), \tag{21}$$

for $\theta \in B_k$. By applying the mean theorem to the first term in the RHS in Eq. (21), we can obtain $\eta_k \in (t_k, t_k + h_k)$, which satisfies $\int_{B_k} r_f'(\theta)^2(h_k/2 + t_k - \theta)^2 d\theta = r_f'(\eta_k)^2 \int_{B_k}(h_k/2 + t_k - \theta)^2 d\theta$. Thus, we obtain:

$$\begin{aligned}\int_{B_k} \mathrm{Bias}[\hat{r}_f(\theta;h)]^2 d\theta &= r'_f(\eta_k)^2 \int_{B_k} (\theta - t_k - h_k/2)^2 d\theta [1 + o(1)] \\ &= r'_f(\eta_k)^2 (h_k^3/12)[1 + o(1)].\end{aligned} \tag{22}$$

Let $\mathrm{ISB} := \int_0^{2\pi} \mathrm{Bias}[\hat{r}_f(\theta;h)]^2 d\theta$. We find that ISB is the sum of the integrations in B_k in (22) for $k = 1, 2, \ldots, m$ with $m := m_n$, such as $m_n \to \infty$ as $n \to \infty$; that is,

$$\begin{aligned}\mathrm{ISB} &= \frac{h^2}{12} \sum_{k=0}^{m-1} r'_f(\eta_k)^2 h[1 + o(1)] + \frac{c^3 h^3}{12} r'_f(\eta_m)^2 [1 + o(1)] \\ &= \frac{h^2}{12} \left[\int_0^{2\pi} r'_f(\theta)^2 d\theta - \int_{2\pi - ch}^{2\pi} r'_f(\theta)^2 d\theta + c^3 h r'_f(\eta_m)^2 \right] [1 + o(1)] \\ &= \frac{h^2}{12} \int_0^{2\pi} r'_f(\theta)^2 d\theta [1 + o(1)],\end{aligned} \tag{23}$$

where we use the Riemann integral approximation in the second RHS and $\int_{2\pi-ch}^{2\pi} r'_f(\theta)^2 d\theta - ch r'_f(\eta_m)^2 = o(1)$ in the last RHS. From Lemma 1 and (17), we obtain:

$$\begin{aligned}\mathrm{Var}[\hat{r}_f(\theta;h)] &= \mathrm{Var}[2^{-1/2} \mathrm{E}[\hat{f}(\theta;h_k)]^{-1/2} W_k + O_p(W_k^2)] \\ &= (2\mathrm{E}[\hat{f}(\theta;h_k)])^{-1} \mathrm{E}[W_k^2] + O(\mathrm{E}[W_k^4]) \\ &= (2f(\xi_k))^{-1} [f(\xi_k)/(nh_k) - f(\xi_k)^2/n] + O((nh)^{-2}) \\ &= 1/(2nh_k) - f(\xi_k)/(2n) + O((nh)^{-2}).\end{aligned} \tag{24}$$

Let $\mathrm{IV} := \int_0^{2\pi} \mathrm{Var}[\hat{r}_f(\theta;h)] d\theta$. Integrating (24) provides:

$$\begin{aligned}\mathrm{IV} &= \int_0^{2\pi} \frac{1}{2nh} d\theta - \int_{2\pi-ch}^{2\pi} \frac{1}{2nh} d\theta + \int_{2\pi-ch}^{2\pi} \frac{1}{2cnh} d\theta + O(n^{-1} + (nh)^{-2}) \\ &= \pi/(nh) + o((nh)^{-1}).\end{aligned} \tag{25}$$

Combining (23) and (25) leads to:

$$\mathrm{MISE}[\hat{r}_f(\theta;h)] = h^2 R(r'_f)/12 + \pi/(nh) + o(h^2 + (nh)^{-1}). \tag{26}$$

Theorem 1 completes the proof from combining (21), (24), and (26). □

Appendix B

We describe the expectation, variance, and covariance of the histogram estimator. We base these proofs on the results in Scott (1979, 1985). We use these to derive the properties of the rose diagram estimator.

The Taylor expansion of f yields:

$$\begin{aligned} p_k &= \int_{t_k}^{t_k+h_k} f(\theta)d\theta \\ &= \int_{t_k}^{t_k+h_k} \left[f(t_k)+f'(t_k)(\theta-t_k)+\frac{1}{2}f''(t_k)(\theta-t_k)^2+O((\theta-t_k)^3)\right]d\theta \\ &= f(t_k)h_k+\frac{1}{2}f'(t_k)h_k^2+\frac{1}{6}f''(t_k)h_k^3+O(h^4). \end{aligned} \tag{27}$$

Note that $p_{k-1}=\int_{t_k-h_{k-1}}^{t_k} f(\theta)d\theta$ and $p_0=\int_{t_{m+1}-h_m}^{t_{m+1}} f(\theta)d\theta$. The Taylor expansion of f also yields:

$$p_{k-1}=\begin{cases} f(t_k)h_{k-1}-\frac{1}{2}f'(t_k)h_{k-1}^2+\frac{1}{6}f''(t_k)h_{k-1}^3+O(h^4) & \text{if} \quad k\in\{2,3,\ldots,m\}, \\ f(t_1)h_m-\frac{1}{2}f'(t_1)h_m^2+\frac{1}{6}f''(t_1)h_m^3+O(h^4), & \text{if} \quad k=1, \end{cases} \tag{28}$$

where use $f(t_{m+1})=f(t_1)$. From (27), we obtain:

$$\mathrm{E}[\hat{f}_k]=f(t_k)+\frac{1}{2}f'(t_k)h_k+\frac{1}{6}f''(t_k)h_k^2+O(h^3), \tag{29}$$

and

$$\begin{aligned} \mathrm{E}[W_k^2] &= \left[f(t_k)h_k+\frac{f'(t_k)h_k^2}{2}+O(h^3)\right]\frac{1-f(t_k)h_k+O(h^2)}{nh_k^2} \\ &= \frac{f(t_k)}{nh_k}+\frac{f'(t_k)}{2n}-\frac{f(t_k)^2}{n}+O(n^{-1}h). \end{aligned} \tag{30}$$

From (28), we obtain:

$$\mathrm{E}[\hat{f}_{k-1}]=f(t_k)-\frac{1}{2}f'(t_k)h_{k-1}+\frac{1}{6}f''(t_k)h_{k-1}^2+O(h^3), \tag{31}$$

and

$$\mathrm{E}[W_{k-1}^2]=\begin{cases} \frac{f(t_k)}{nh_{k-1}}-\frac{f'(t_k)}{2n}+\frac{f(t_k)^2}{n}+O(n^{-1}h) & \text{if} \quad k\in\{2,3,\ldots,m\}, \\ \frac{f(t_1)}{nh_m}-\frac{f'(t_1)}{2n}+\frac{f(t_1)^2}{n}+O(n^{-1}h) & \text{if} \quad k=1. \end{cases} \tag{32}$$

We know that polynomial distribution's covariance is $\mathrm{Cov}[v_k,v_{k-1}]=-np_kp_{k-1}$. Thus, combining (27) and (28) yields

$$\begin{aligned} \mathrm{E}[W_kW_{k-1}] &= -\frac{np_kp_{k-1}}{n^2h_kh_{k-1}} \\ &= -\frac{[f(t_k)h_k+O(h^2)][f(t_k)h_{k-1}+O(h^2)]}{nh_kh_{k-1}} \\ &= -\frac{f(t_k)^2}{n}+O(n^{-1}h), \end{aligned} \tag{33}$$

where use $h_0 = h_m$.

Appendix C

Proof of Theorem 2 First, we derive the expectation $\hat{R}_1$. Let the histogram estimator $\hat{f}(\cdot)$ in bin B_k be $\hat{f}_k$. We obtain:

$$\mathrm{E}[\hat{R}_1] = 2\sum_{k=1}^{m} \tilde{h}_k^{-1}(\mathrm{E}[\hat{f}_k] + \mathrm{E}[\hat{f}_{k-1}] - \mathrm{E}[\hat{r}_k \hat{r}_{k-1}]). \tag{34}$$

We obtain the following Taylor expansion of $\hat{r}_k \hat{r}_{k-1}$:

$$\begin{aligned}\hat{r}_k \hat{r}_{k-1} = {} & 2(\mathrm{E}[\hat{f}_k]\mathrm{E}[\hat{f}_{k-1}])^{1/2} + \mathrm{E}[\hat{f}_k]^{-1/2}\mathrm{E}[\hat{f}_{k-1}]^{1/2} W_k + \mathrm{E}[\hat{f}_k]^{1/2}\mathrm{E}[\hat{f}_{k-1}]^{-1/2} W_{k-1} \\ & - \frac{1}{4}\mathrm{E}[\hat{f}_k]^{-3/2}\mathrm{E}[\hat{f}_{k-1}]^{1/2} W_k^2 - \frac{1}{4}\mathrm{E}[\hat{f}_k]^{1/2}\mathrm{E}[\hat{f}_{k-1}]^{-3/2} W_{k-1}^2 \\ & + \frac{1}{2}\mathrm{E}[\hat{f}_k]^{-1/2}\mathrm{E}[\hat{f}_{k-1}]^{-1/2} W_k W_{k-1} + O_p((W_k + W_{k-1})^3 + (W_k + W_{k-1})^4).\end{aligned} \tag{35}$$

Combining Lemma 1 and (35) yields $\mathrm{E}[\hat{r}_k \hat{r}_{k-1}]$, which is equal to

$$\begin{aligned}\mathrm{E}[\hat{r}_k \hat{r}_{k-1}] = {} & 2(\mathrm{E}[\hat{f}_k]\mathrm{E}[\hat{f}_{k-1}])^{1/2} - \frac{1}{4}\mathrm{E}[\hat{f}_k]^{-3/2}\mathrm{E}[\hat{f}_{k-1}]^{1/2}\mathrm{E}[W_k^2] \\ & - \frac{1}{4}\mathrm{E}[\hat{f}_k]^{1/2}\mathrm{E}[\hat{f}_{k-1}]^{-3/2}\mathrm{E}[W_{k-1}^2] + \frac{1}{2}\mathrm{E}[\hat{f}_k]^{-1/2}\mathrm{E}[\hat{f}_{k-1}]^{-1/2}\mathrm{E}[W_k W_{k-1}] \\ & + O((nh)^{-2}).\end{aligned} \tag{36}$$

Combining (34) and (36) provides:

$$\begin{aligned}\mathrm{E}[\hat{R}_1] = 2\sum_{k=1}^{m} \tilde{h}_k^{-1} \Big[& (\mathrm{E}[\hat{f}_k]^{1/2} - \mathrm{E}[\hat{f}_{k-1}]^{1/2})^2 + \frac{1}{4}\mathrm{E}[\hat{f}_k]^{-3/2}\mathrm{E}[\hat{f}_{k-1}]^{1/2}\mathrm{E}[W_k^2] \\ & + \frac{1}{4}\mathrm{E}[\hat{f}_k]^{1/2}\mathrm{E}[\hat{f}_{k-1}]^{-3/2}\mathrm{E}[W_{k-1}^2] - \frac{1}{2}\mathrm{E}[\hat{f}_k]^{-1/2}\mathrm{E}[\hat{f}_{k-1}]^{-1/2}\mathrm{E}[W_k W_{k-1}] \\ & + O((nh)^{-2})\Big].\end{aligned} \tag{37}$$

We now derive each term in the RHS in Eq. (37). Combining (29) and (31) (in Appendix B) leads to:

$$
\begin{aligned}
&(\mathrm{E}[\hat{f}_k]^{1/2} - \mathrm{E}[\hat{f}_{k-1}]^{1/2})^2 \\
&= \left\{ \left[f(t_k) + \frac{1}{2} f'(t_k) h_k + O(h^2) \right]^{1/2} - \left[f(t_k) - \frac{1}{2} f'(t_k) h_{k-1} + O(h^2) \right]^{1/2} \right\}^2 \\
&= \left\{ f(t_k)^{1/2} + \frac{1}{4} f(t_k)^{-1/2} f'(t_k) h_k - \left[f(t_k)^{1/2} - \frac{1}{4} f(t_k)^{-1/2} f'(t_k) h_{k-1} \right] + O(h^2) \right\}^2 \\
&= \left[\frac{1}{2} f(t_k)^{-1/2} f'(t_k) \tilde{h}_k + O(h^2) \right]^2 \\
&= \frac{1}{4} f(t_k)^{-1} f'(t_k)^2 \tilde{h}_k^2 + O(h^3) \\
&= \frac{1}{2} r_f'(t_k)^2 \tilde{h}_k^2 + O(h^3).
\end{aligned}
\tag{38}
$$

Combining (29), (30), and (31) provides:

$$
\begin{aligned}
\mathrm{E}[\hat{f}_k]^{-3/2} \mathrm{E}[\hat{f}_{k-1}]^{1/2} \mathrm{E}[W_k^2] &= \frac{[f(t_k) + O(h)]^{1/2}}{[f(t_k) + O(h)]^{3/2}} \left[\frac{f(t_k)}{n h_k} + O(n^{-1}) \right] \\
&= \frac{1}{n h_k} + O(n^{-1}).
\end{aligned}
\tag{39}
$$

Combining (29), (31), and (32) provides:

$$
\mathrm{E}[\hat{f}_k]^{1/2} \mathrm{E}[\hat{f}_{k-1}]^{-3/2} \mathrm{E}[W_{k-1}^2] = \frac{1}{n h_{k-1}} + O(n^{-1}). \tag{40}
$$

Combining (29), (31), and (33) provides:

$$
\mathrm{E}[\hat{f}_k]^{-1/2} \mathrm{E}[\hat{f}_{k-1}]^{-1/2} \mathrm{E}[W_k W_{k-1}] = -\frac{f(t_k)}{n} + O(n^{-1} h). \tag{41}
$$

By combining (37), (38), (39), (40), and (41), we obtain:

$$
\begin{aligned}
\mathrm{E}[\hat{R}_1] &= 2 \sum_{k=1}^{m} \tilde{h}_k^{-1} \left[\frac{1}{2} r_f'(t_k)^2 \tilde{h}_k^2 + \frac{1}{4 n h_k} + \frac{1}{4 n h_{k-1}} + O(h^3 + n^{-1} + (nh)^{-2}) \right] \\
&= \left[\sum_{k=1}^{m} r_f'(t_k)^2 \tilde{h}_k + \sum_{k=1}^{m} \frac{1}{2 n h_k \tilde{h}_k} + \sum_{k=1}^{m} \frac{1}{2 n h_{k-1} \tilde{h}_k} \right] [1 + o(1)] \\
&= \left[\int_0^{2\pi} r_f'(\theta)^2 d\theta + \sum_{k=2}^{m-1} \frac{1}{nh^2} + \frac{2}{n(h^2 + ch^2)} + \frac{2}{n(ch^2 + c^2 h^2)} \right] [1 + o(1)] \\
&= R(r_f')[1 + o(1)] + \left[\frac{2\pi - (1 + c)h}{nh^3} + \frac{2}{n(h^2 + ch^2)} + \frac{2}{n(ch^2 + c^2 h^2)} \right] [1 + o(1)] \\
&= R(r_f') + \frac{2\pi}{nh^3} + o(1 + n^{-1} h^{-3}).
\end{aligned}
\tag{42}
$$

From (42), we obtain:

$$\begin{aligned}\mathrm{E}[\mathrm{BCV}(h)] &= \frac{h^2}{12}\left\{\mathrm{E}[\hat{R}_1] - \frac{2\pi}{nh^3}\right\} + \frac{\pi}{nh} \\ &= \frac{h^2 R(r_f')}{12} + \frac{\pi}{nh} + o(h^2 + (nh)^{-1}) \\ &= \mathrm{AMISE}[\hat{r}_f(\theta;h)] + o(h^2 + (nh)^{-1}).\end{aligned} \tag{43}$$

We reduce $\hat{R}_1$ to the simpler form. Recalling that $B_0 = B_m$ implies $v_0 = v_m$ and $h_0 = h_m$, we obtain:

$$\begin{aligned}\hat{R}_1 &= \sum_{k=1}^{m}(2\hat{f}_k + 2\hat{f}_{k-1} - 2\hat{r}_k\hat{r}_{k-1})\tilde{h}_k^{-1} \\ &= 2\left[\sum_{k=2}^{m-1}\frac{v_k}{nh^2} + \sum_{k=2}^{m-1}\frac{v_{k-1}}{nh^2} + \frac{v_1}{nh\tilde{h}_1} + \frac{v_m}{nh_m\tilde{h}_m} + \frac{v_0}{nh_0\tilde{h}_1} - \frac{v_{m-1}}{nh\tilde{h}_m} - \sum_{k=1}^{m}\tilde{h}_k^{-1}\hat{r}_k\hat{r}_{k-1}\right] \\ &= 2\Bigg\{2\sum_{k=1}^{m}\frac{v_k}{nh^2} + \left[\frac{2}{(1+c)h} - \frac{1}{h}\right](\hat{f}_1 + \hat{f}_{m-1}) + 2\left[\frac{2}{c(1+c)h} - \frac{1}{h}\right]\hat{f}_m \\ &\quad - \sum_{k=1}^{m}\tilde{h}_k^{-1}\hat{r}_k\hat{r}_{k-1}\Bigg\} \\ &= \frac{4}{h^2} + \frac{2a(c)}{h}(\hat{f}_1 + \hat{f}_{m-1}) + \frac{4b(c)}{h}\hat{f}_m - 2\sum_{k=1}^{m}\tilde{h}_k^{-1}\hat{r}_k\hat{r}_{k-1},\end{aligned} \tag{44}$$

where $a(c) := (1-c)/(1+c)$ and $b(c) := [2-(c+c^2)]/(c+c^2)$. Next, we provide the variance of $\mathrm{BCV}(h)$. By combining (11) and (44), we reduce $\mathrm{BCV}(h)$ to

$$\mathrm{BCV}(h) = \frac{1}{3} + \frac{5\pi}{6nh} + \frac{ha(c)}{6}(\hat{f}_1 + \hat{f}_{m-1}) + \frac{hb(c)}{3}\hat{f}_m - \frac{h^2}{6}\sum_{k=1}^{m}\tilde{h}_k^{-1}\hat{r}_k\hat{r}_{k-1}. \tag{45}$$

Let $Y_{k,k-1} := \hat{r}_k\hat{r}_{k-1} - \mathrm{E}[\hat{r}_k\hat{r}_{k-1}]$. Then, from (45), we obtain:

$$\begin{aligned}\mathrm{Var}[\mathrm{BCV}(h)] &= \frac{h^4}{36}\sum_{k=1}^{m}\tilde{h}_k^{-2}\mathrm{E}[Y_{k,k-1}^2] + \frac{h^2a(c)^2}{36}(\mathrm{E}[W_1^2] + \mathrm{E}[W_{m-1}^2]) + \frac{h^2b(c)^2}{9}\mathrm{E}[W_m^2] \\ &\quad - \frac{h^3a(c)}{18}\sum_{k=1}^{m}\tilde{h}_k^{-1}(\mathrm{E}[W_1Y_{k,k-1}] + \mathrm{E}[W_{m-1}Y_{k,k-1}]) - \frac{h^3b(c)}{9}\sum_{k=1}^{m}\tilde{h}_k^{-1}\mathrm{E}[W_mY_{k,k-1}] \\ &\quad + \frac{h^4}{18}\sum_{k<l}\tilde{h}_k^{-1}\tilde{h}_l^{-1}\mathrm{E}[Y_{k,k-1}Y_{l,l-1}] + \frac{h^2a(c)^2}{18}\mathrm{E}[W_1W_{m-1}] \\ &\quad + \frac{h^2a(c)b(c)}{9}(\mathrm{E}[W_1W_m] + \mathrm{E}[W_{m-1}W_m]).\end{aligned} \tag{46}$$

Combining (35) and (36) provides the following approximations of $Y_{k,k-1}$:

$$
\begin{aligned}
Y_{k,k-1} &= \mathrm{E}[\hat{f}_k]^{-1/2}\mathrm{E}[\hat{f}_{k-1}]^{1/2}W_k + \mathrm{E}[\hat{f}_k]^{1/2}\mathrm{E}[\hat{f}_{k-1}]^{-1/2}W_{k-1} \\
&\quad - \frac{1}{4}\mathrm{E}[\hat{f}_k]^{-3/2}\mathrm{E}[\hat{f}_{k-1}]^{1/2}(W_k^2 - \mathrm{E}[W_k^2]) - \frac{1}{4}\mathrm{E}[\hat{f}_k]^{1/2}\mathrm{E}[\hat{f}_{k-1}]^{-3/2}(W_{k-1}^2 - \mathrm{E}[W_{k-1}^2]) \\
&\quad + \frac{1}{2}\mathrm{E}[\hat{f}_k]^{-1/2}\mathrm{E}[\hat{f}_{k-1}]^{-1/2}(W_kW_{k-1} - \mathrm{E}[W_kW_{k-1}]) \\
&\quad + O_p((W_k + W_{k-1})^3 + (W_k + W_{k-1})^4) + O((nh)^{-2}).
\end{aligned} \tag{47}
$$

By combining, (29), (30), (31), (32) (33), and (47) we obtain the following simple form:

$$
Y_{k,k-1} = [1 + O(h)](W_k + W_{k-1}) + O_p((W_k + W_{k-1})^2) + O((nh)^{-1}). \tag{48}
$$

Combining, (30), (32) (33), and (48) yields:

$$
\begin{aligned}
\mathrm{E}[Y_{k,k-1}^2] &= [1 + O(h)]^2(\mathrm{E}[W_k^2] + \mathrm{E}[W_{k-1}^2] + \mathrm{E}[W_kW_{k-1}]) + O((nh)^{-2}) \\
&= \frac{f(t_k)}{nh_k} + \frac{f(t_k)}{nh_{k-1}} + O(n^{-1} + (nh)^{-2}),
\end{aligned} \tag{49}
$$

where $h_0 = h_m$. Using the same method as (33), we show that $\mathrm{E}[W_kW_l] = O(n^{-1})$ for any $k \neq l$. Therefore, combining this, (30), and (48) yields the following covariances:

$$
\begin{aligned}
\mathrm{E}[W_jY_{k,k-1}] &= \mathrm{E}\big[[1 + O(h)](W_jW_k + W_jW_{k-1}) + O_p(W_j(W_k + W_{k-1})^2)\big] \\
&= [1 + O(h)](\mathrm{E}[W_jW_k] + \mathrm{E}[W_jW_{k-1}]) \\
&= \begin{cases} f(t_j)(nh_j)^{-1} + O(n^{-1}) & \text{if } k = j \text{ or } k = j+1, \\ O(n^{-1}) & \text{otherwise,} \end{cases}
\end{aligned} \tag{50}
$$

for $j \in \{1, 2, \ldots, m\}$,

$$
\mathrm{E}[W_mY_{k,k-1}] = \begin{cases} f(t_m)(nh_m)^{-1} + O(n^{-1}) & \text{if } k = m \text{ or } k = 1, \\ O(n^{-1}) & \text{otherwise,} \end{cases} \tag{51}
$$

and

$$
\begin{aligned}
\mathrm{E}[Y_{k,k-1}, Y_{l,l-1}] &= [1 + O(h)]^2(\mathrm{E}[W_kW_l] + \mathrm{E}[W_kW_{l-1}] + \mathrm{E}[W_{k-1}W_l] + \mathrm{E}[W_{k-1}W_{l-1}]) \\
&\quad + O((nh)^{-2}) \\
&= O(n^{-1} + (nh)^{-2}).
\end{aligned} \tag{52}
$$

Combining (33) and (52) demonstrates that the magnitude of the sixth, seventh, and eighth terms in the RHS in (46) is $O(n^{-1}h^2)$. This indicates that these terms are ignored. Therefore, by combining (30), (32), (46), (49), (50), and (51), we obtain:

$$
\begin{aligned}
\mathrm{Var}[\mathrm{BCV}(h)] &= \frac{h^4}{36}\left[2\sum_{k=2}^{m-1}\frac{f(t_k)}{nh^3} + \frac{f(t_1)}{nh\tilde{h}_1^2} + \frac{f(t_m)}{nh_m\tilde{h}_m^2} + \frac{f(t_m)}{nh\tilde{h}_m^2} + \frac{f(t_1)}{nh^2\tilde{h}_m} + O(n^{-1} + (nh)^{-2})\right] \\
&\quad + \frac{h^2a(c)^2}{36}\left[\frac{f(t_1)}{nh} + \frac{f(t_{m-1})}{nh}\right] + \frac{h^2b(c)^2}{9}\frac{f(t_m)}{nh_m} \\
&\quad - \frac{h^2a(c)}{18}\left[2\frac{f(t_1)}{nh} + \frac{f(t_{m-1})}{nh} + \frac{f(t_{m-1})}{nch}\right] \\
&\quad - \frac{h^3b(c)}{9}\left[\frac{f(t_m)}{nh_m\tilde{h}_m} + \frac{f(t_m)}{nh\tilde{h}_m}\right] + O(n^{-1}h^2) \\
&= \frac{1}{18n}\int_h^{2\pi-ch} f(\theta)d\theta[1 + o(1)] \\
&= \frac{1}{18n} + o(n^{-1}).
\end{aligned}
\tag{53}
$$

Theorem 2 completes the proof from (43) and (53). □

Appendix D

Proof of Corollary 2 We denote the MISE and the asymptotic MISE of the rose diagram estimator as MISE(h) and AMISE(h), respectively. From (5), we obtain:

$$\mathrm{AMISE}(h)/\mathrm{MISE}(h) \xrightarrow{p} 1. \tag{54}$$

From combining Theorem 2 and the law of large numbers, we obtain:

$$\mathrm{BCV}(h)/\mathrm{AMISE}(h) \xrightarrow{p} 1. \tag{55}$$

Let $h_* = h_{\mathrm{opt}}$ and $\gamma := \hat{h}_{\mathrm{BCV}}/h_*$ We find that (54) and (55) hold for all h. Therefore, we obtain:

$$\mathrm{AMISE}(\gamma h_*)/\mathrm{MISE}(\gamma h_*) \xrightarrow{p} 1, \tag{56}$$

and

$$\mathrm{BCV}(\gamma h_*)/\mathrm{MISE}(\gamma h_*) \xrightarrow{p} 1. \tag{57}$$

The ratio of the AMISE of bin width γh_* to that of bin width h_* is a function expressed in the following equation:

$$\mathrm{AMISE}(\gamma h_*)/\mathrm{AMISE}(h_*) = \frac{\gamma^2}{3} + \frac{2}{3\gamma}. \tag{58}$$

The ratio in (58) is a convex function with minimum value at $\gamma = 1$.Thus, if $\gamma \neq 1$ and n is sufficiently large, from (56), we obtain:

$$\mathrm{MISE}(\gamma h_*) > \mathrm{MISE}(h_*). \tag{59}$$

Suppose that γ does not converge to 1.If n is sufficiently large, then $\mathrm{BCV}(h)$ is a convex function with a minimum at $h = \gamma h_*$, as $\mathrm{BCV}(h)$ closes to $\mathrm{AMISE}(h)$. Therefore, we obtain:

$$\mathrm{P}(\mathrm{BCV}(\gamma h_*) < \mathrm{BCV}(h_*)) \xrightarrow{p} 1, \tag{60}$$

as $n \to \infty$. We combine (57) and (60), and find that

$$\mathrm{MISE}(\gamma h_*) < \mathrm{MISE}(h_*), \tag{61}$$

as $n \to \infty$. The contradiction between (59) and (61) completes the proof. □

Appendix E

Proof of Theorem 3 We now calculate the bias. Note that $\hat{r}_k(\theta;h) = \hat{r}_k$. Combining (18) and (29) provides the approximation of $\mathrm{E}[\hat{r}_k]$:

$$\begin{aligned}
\mathrm{E}[\hat{r}_k] &= \sqrt{2}\left(f(t_k) + \frac{1}{2}f'(t_k)h_k + \frac{1}{6}f''(t_k)h_k^2 + O(h^3)\right)^{1/2} + O((nh)^{-1}) \\
&= \sqrt{2f(t_k)} + \frac{f'(t_k)h_k}{2\sqrt{2f(t_k)}} + \frac{f''(t_k)h_k^2}{6\sqrt{2f(t_k)}} - \frac{f'(t_k)^2h_k^2}{16\sqrt{2}f(t_k)^{3/2}} + O(h^3 + (nh)^{-1}) \\
&= r_f(t_k) + \frac{1}{2}r_f'(t_k)h_k + \frac{4r_f''(t_k) + r_f'(t_k)^2 r_f(t_k)^{-1}}{24}h_k^2 + O(h^3 + (nh)^{-1}).
\end{aligned} \tag{62}$$

Similarly, combining (18) and (31) provides the approximation of $\mathrm{E}[\hat{r}_{k-1}]$:

$$\mathrm{E}[\hat{r}_{k-1}] = r_f(t_k) - \frac{1}{2}r_f'(t_k)h_{k-1} + \frac{4r_f''(t_k) + r_f'(t_k)^2 r_f(t_k)^{-1}}{24}h_{k-1}^2 + O(h^3 + (nh)^{-1}). \tag{63}$$

By combining (62) and (63), we obtain:

$$\begin{aligned}
\mathrm{E}[\tilde{r}_f(\theta;h)] = r_f(t_k) &+ r_f'(t_k)(\theta - t_k) \\
&+ \frac{4r_f''(t_k) + r_f'(t_k)^2 r_f(t_k)^{-1}}{24}[h_k h_{k-1} + 2(\theta - t_k)(h_k - h_{k-1})] \\
&+ O(h^3 + (nh)^{-1}),
\end{aligned} \tag{64}$$

for $\theta \in [t_k - h_{k-1}/2, t_k + h_k/2)$. The Taylor expansion of $r_f(\theta)$ for $\theta \in [t_k - h_{k-1}/2, t_k + h_k/2)$ is

$$r_f(\theta) = r_f(t_k) + r'(t_k)(\theta - t_k) + \frac{1}{2}r_f''(t_k)(\theta - t_k)^2 + O((\theta - t_k)^3). \tag{65}$$

Combining (64) and (65) leads to:

$$\begin{aligned}\mathrm{Bias}[\tilde{r}(\theta;h)] &= \frac{r_f''(t_k)}{6}[h_k h_{k-1} + 2(\theta - t_k)(h_k - h_{k-1}) - 3(\theta - t_k)^2] \\ &+ \frac{1}{24} r_f'(t_k)^2 r_f(t_k)^{-1}[h_k h_{k-1} + 2(\theta - t_k)(h_k - h_{k-1})] \\ &+ O(h^3 + (nh)^{-1}).\end{aligned} \tag{66}$$

Let $\mathrm{ISB} := \int_0^{2\pi} \mathrm{Bias}[\tilde{r}(\theta;h)]^2 d\theta$. Note that $\mathrm{ISB} = \int_{h/2}^{2\pi-(c+1/2)h} \mathrm{Bias}[\tilde{r}(\theta;h)]^2 d\theta + o(1) = \sum_{k=2}^{m-1} \int_{t_k-h/2}^{t_k+h/2} \mathrm{Bias}[\tilde{r}(\theta;h)]^2 d\theta [1 + o(1)]$. For $\theta \in [t_k - h/2, t_k + h/2)$, $k = 2, 3, \ldots, m-1$, $\mathrm{Bias}[\tilde{r}(\theta;h)]$ in (66) is

$$\mathrm{Bias}[\tilde{r}(\theta;h)] = \frac{r_f''(t_k)}{6}[h^2 - 3(\theta - t_k)^2] + \frac{1}{24} r_f'(t_k)^2 r_f(t_k)^{-1} h^2 + O(h^3 + (nh)^{-1}). \tag{67}$$

Integration of $\mathrm{Bias}[\tilde{r}(\theta;h)]$ in (67) for $\theta \in [t_k - h/2, t_k + h/2)$, $k = 2, 3, \ldots, m-1$ is

$$\begin{aligned}&\int_{t_k-h/2}^{t_k+h/2} \mathrm{Bias}[\tilde{r}(\theta;h)]^2 d\theta \\ &= \int_{t_k-h/2}^{t_k+h/2} \left\{ \frac{r_f''(t_k)}{36}[h^4 - 6(\theta - t_k)^2 + 9(\theta - t_k)^4] + \frac{1}{72} r_f''(t_k) r_f'(t_k)^2 r_f(t_k)^{-1}[h^4 - 3(\theta - t_k)^2 h^2] \right. \\ &\quad \left. + \frac{1}{576} r_f'(t_k)^4 r_f(t_k)^{-2} h^4 \right\} d\theta [1 + o(1)] \\ &= \left[\frac{49}{2880} r_f''(t_k)^2 + \frac{1}{96} r_f''(t_k) r_f'(t_k)^2 r_f(t_k)^{-1} + \frac{1}{576} r_f'(t_k)^4 r_f(t_k)^{-2} \right] h^5 [1 + o(1)].\end{aligned} \tag{68}$$

We find that $\sum_{k=2}^{m-1} r_f''(t_k)^2 h = \int_{h/2}^{2\pi-(c+1/2)h} r_f''(\theta)^2 d\theta + o(1) = \int_0^{2\pi} r_f''(\theta)^2 d\theta + \mathrm{o}(1)$. Calculating each term in (68) for $k = 2, 3, \ldots, m-1$ yields the following form:

$$\mathrm{ISB} = \left\{ \frac{49}{2880} R(r_f'') + \frac{1}{96} R((r_f'')^{1/2} r_f' r_f^{-1/2}) + \frac{1}{576} R((r_f')^2 r_f^{-1}) \right\} h^4 + o(h^4). \tag{69}$$

Next, we derive $\mathrm{IV} := \int_0^{2\pi} \mathrm{Var}[\tilde{r}(\theta;h)] d\theta$. We calculate the variances of $\hat{r}_k$ and $\hat{r}_{k-1}$ by applying the same method as in (24). By combining (17), (29), (30), (31), and (32), we obtain:

$$\begin{aligned}\mathrm{Var}[\hat{r}_k] &= (2\mathrm{E}[\hat{f}_k])^{-1} \mathrm{E}[W_k^2] + O((nh)^{-2}) \\ &= \frac{1}{2nh_k} + O(n^{-1} + (nh)^{-2}),\end{aligned} \tag{70}$$

and

$$\mathrm{Var}[\hat{r}_{k-1}] = \frac{1}{2nh_{k-1}} + O(n^{-1} + (nh)^{-2}). \tag{71}$$

Combining (17), (29), (31), and (33) leads to:

$$\begin{aligned}\mathrm{Cov}[\hat{r}_{k-1},\hat{r}_k] &= 2^{-1}(E[\hat{f}_k]E[\hat{f}_{k-1}])^{-1/2}\mathrm{E}[W_kW_{k-1}] + O((nh)^{-2}) \\ &= \frac{1}{2[f(t_k)+O(h)]}\left[-\frac{f(t_k)^2}{n} + O(n^{-1}h)\right] + O((nh)^{-2}) \\ &= -\frac{f(t_k)}{2n} + O(n^{-1}h + (nh)^{-2}).\end{aligned} \tag{72}$$

By combining (70), (71), and (72), we obtain:

$$\mathrm{Var}[\tilde{r}_f(\theta;h)] = \left(\frac{h_k}{2\tilde{h}_k} - \frac{\theta - t_k}{\tilde{h}_k}\right)^2 \frac{1}{2nh_{k-1}} + \left(\frac{h_{k-1}}{2\tilde{h}_k} + \frac{\theta - t_k}{\tilde{h}_k}\right)^2 \frac{1}{2nh_k} + o((nh)^{-1}). \tag{73}$$

Note that $\mathrm{IV} = \sum_{k=2}^{m-1}\int_{t_k-h/2}^{t_k+h/2}\mathrm{Var}[\tilde{r}(\theta;h)]d\theta[1+o(1)]$. For $\theta \in [t_k - h/2, t_k + h/2)$, $k = 2, 3, \ldots, m-1$, from (73), we obtain:

$$\mathrm{Var}[\tilde{r}_f(\theta;h)] = \frac{1}{2nh}\left[\frac{1}{2} + 2\frac{(\theta - t_k)^2}{h^2}\right] + o((nh)^{-1}). \tag{74}$$

Integration of $\mathrm{Var}[\tilde{r}_f(\theta;h)]$ in (74) for $\theta \in [t_k - h/2, t_k + h/2)$, $k = 2, 3, \ldots, m-1$ is

$$\int_{t_k-h/2}^{t_k+h/2}\mathrm{Var}[\tilde{r}(\theta;h)]d\theta = \frac{1}{3n}[1+o(1)]. \tag{75}$$

From (75), we obtain:

$$\begin{aligned}\mathrm{IV} &= \sum_{k=2}^{m-1}\int_{t_k-h/2}^{t_k+h/2}\mathrm{Var}[\tilde{r}(\theta;h)]^2 d\theta[1+o(1)] \\ &= \sum_{k=2}^{m-1}\frac{1}{3nh}[1+o(1)]h \\ &= \frac{2\pi - (1+c)h}{3nh} + o((nh)^{-1}) \\ &= \frac{2\pi}{3nh} + o((nh)^{-1}).\end{aligned} \tag{76}$$

Theorem 3 completes the proof from combining (69) and (76). □

References

Abe, T., Pewsey, A. (2011). Sine-skewed circular distributions. *Statistical Papers, 52*, 683–707.

Agostinelli, C., Lund, U. (2017). *R package* `circular`: *Circular Statistics* (*version 0.4-93*), CA: Department of Environmental Sciences, Informatics and Statistics, Ca' Foscari University, Venice, Italy. UL: Department of Statistics, California Polytechnic State University, San Luis Obispo, California, USA, https://r-forge.r-project.org/projects/circular/.

Andreasen, C. (1990). A suggested standard procedure for the construction of unimodal current rose-diagrams. *Journal of Sedimentary Research, 60*(4), 628–629.

Ballantyne, C., Cornish, R. (1979). Use of the chi-square test for the analysis of orientation data. *Journal of Sedimentary Research, 49*(3), 773–776.
Di Marzio, M., Panzera, A., Taylor, C. C. (2011). Kernel density estimation on the torus. *Journal of Statistical Planning and Inference, 141*(6), 2156–2173.
Fisher, N. I. (1993). *Statistical analysis of circular data*. New York: Cambridge University Press.
Hall, P., Watson, G., Cabrera, J. (1987). Kernel density estimation with spherical data. *Biometrika, 74*(4), 751–762.
Kanazawa, Y. (1993). Hellinger distance and Akaike's information criterion for the histogram. *Statistics and Probability Letters, 17*(4), 293–298.
Mardia, K. V. (1972). *Statistics of directional data*. London: Academic Press.
Mardia, K. V., Jupp, P. E. (2000). *Directional statistics*. New York: Wiley.
Munro, M. A., Blenkinsop, T. G. (2012). MARD–A moving average rose diagram application for the geosciences. *Computers and Geosciences, 49*, 112–120.
Nemec, W. (1988). The shape of the rose. *Sedimentary Geology, 59*(1–2), 149–152.
Nightingale, F. (1858). *Notes on matters affecting the health, efficiency, and hospital administration of the British army*. London: Harrison & Sons.
Pewsey, A., Neuhäuser, M., Ruxton, G. D. (2013). *Circular statistics in R*. Oxford: Oxford University Press.
R Core Team (2021). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/.
Sanderson, D. J., Peacock, D. C. (2020). Making rose diagrams fit-for-purpose. *Earth-Science Reviews, 201*, 103055.
Scott, D. W. (1979). On optimal and data-based histograms. *Biometrika, 66*(3), 605–610.
Scott, D. W. (1985). Frequency polygons: Theory and application. *Journal of the American Statistical Association, 80*(390), 348–354.
Scott, D. W. (2015). *Multivariate density estimation*: *Theory, practice, and visualization*. Hoboken, New Jersey: Wiley.
Scott, D. W., Terrell, G. R. (1987). Biased and unbiased cross-validation in density estimation. *Journal of the American Statistical Association, 82*(400), 1131–1146.
Taylor, C. C. (2008). Automatic bandwidth selection for circular density estimation. *Computational Statistics and Data Analysis, 52*(7), 3493–3500.
Tsuruta, Y., Sagae, M. (2017). Higher order kernel density estimation on the circle. *Statistics and Probability Letters, 131*, 46–50.
Tsuruta, Y., Sagae, M. (2020). Theoretical properties of bandwidth selectors for kernel density estimation on the circle. *Annals of the Institute of Statistical Mathematics, 72*(2), 511–530.
Wells, N. A. (2000). Are there better alternatives to standard rose diagrams? *Journal of Sedimentary Research, 70*(1), 37–46.