

Document Version

Final published version

Licence

Dutch Copyright Act (Article 25fa)

Citation (APA)

Ryan, V., & Derumigny, A. (2025). On the Choice of the Two Tuning Parameters for Nonparametric Estimation of an Elliptical Distribution Generator. In A. Steland, E. Rafajlowicz, & N. Parolya (Eds.), *Stochastic Models, Statistics and Their Applications, SMSA 2024* (pp. 81-124). (Springer Proceedings in Mathematics and Statistics; Vol. 499). Springer. https://doi.org/10.1007/978-3-031-96015-4_5

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership. Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

On the Choice of the Two Tuning Parameters for Nonparametric Estimation of an Elliptical Distribution Generator



Victor Ryan and Alexis Derumigny 

Abstract Elliptical distributions are a simple and flexible class of distributions that depend on a one-dimensional function, called the density generator. In this chapter, we study the nonparametric estimator of this generator that was introduced by Liebscher (J Multivariate Anal 92(1):205–225, 2005). This estimator depends on two tuning parameters: a bandwidth h —as usual in kernel smoothing—and an additional parameter a that control the behavior near the center of the distribution. We give an explicit expression for the asymptotic MSE at a point x and derive explicit expressions for the optimal tuning parameters h and a . The estimation of the derivatives of the generator is also discussed. A simulation study shows the performance of the new methods.

Keywords Elliptical distribution · Kernel smoothing · Optimal bandwidth

1 Introduction

Elliptical distributions are a well-known family of distributions, generalizing spherically symmetric distributions (which are invariant by rotations). In many situations, elliptical distributions allow easy computations such as the Value-at-Risk in finance, see, e.g., [14]. They also facilitate the inference and, in some sense, avoid the curse of dimensionality. Indeed, a continuous d -dimensional elliptical distribution is parametrized by its mean $\mu \in \mathbb{R}^d$, its matrix Σ assumed to be positive definite of dimension $d \times d$, and its density generator $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$. We say that a random vector \mathbf{X} follows the elliptical distribution $\mathcal{E}(\mu, \Sigma, g)$ if its density is

$$f_{\mathbf{X}}(\mathbf{x}) = |\Sigma|^{-1/2} g \left((\mathbf{x} - \mu)^\top \Sigma^{-1} (\mathbf{x} - \mu) \right), \text{ for any } \mathbf{x} \in \mathbb{R}^d. \quad (1)$$

V. Ryan · A. Derumigny (✉)

Department of Applied Mathematics, Delft University of Technology, Delft, Netherlands

e-mail: a.f.derumigny@tudelft.nl

This construction reduces the complexity of the density estimation problem by reducing the estimation of a d -dimensional density function $f_{\mathbf{X}}$ to the estimation of the generator g , a one-dimensional function.

The estimation of the generator g has already been studied, first by [15] who proposed a kernel-based estimator. This estimator was then improved by [11]. An alternative estimator based on sieves was proposed by [1]. Estimators of g based on nonparametric maximum likelihood estimation via isotonic regression or splines were proposed by [2] under the assumption that g is monotone. On a related area, elliptical distributions are the cornerstone of elliptical copula models, which also relies on a generator that needs to be estimated. Several properties of this model, including identifiability conditions, and an estimation algorithm have been studied in [4]. In a Bayesian framework, the work [10] proposes to estimate the generator of an elliptical copula using B-splines.

In this chapter, we are interested in nonparametric estimation of g in a general context, assuming only some smoothness conditions on g . For this, we will rely on Liebscher's kernel-based estimator [11] defined for $\xi \in \mathbb{R}_+$ by

$$\widehat{g}_{n,h,a}(\xi) := \frac{\xi^{-\frac{d+2}{2}} \psi'_a(\xi)}{nhs_d} \sum_{i=1}^n \left[K \left(\frac{\psi_a(\xi) - \psi_a(\xi_i)}{h} \right) + K \left(\frac{\psi_a(\xi) + \psi_a(\xi_i)}{h} \right) \right], \quad (2)$$

where:

- $\mathbf{X}_1, \dots, \mathbf{X}_n$ is an i.i.d. sample from the law $\mathcal{E}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$.
- $\xi_i := (\mathbf{X}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{X}_i - \boldsymbol{\mu})$.
- $h = h(n) > 0$ is the bandwidth, and K is a one-dimensional kernel with compact support such that $K(0) = 1$; K is symmetric and $\int K = 1$.
- $a > 0$ is a second tuning parameter and $\psi_a(\xi) := -a + (a^{d/2} + \xi^{d/2})^{2/d}$.
- $s_d := \pi^{d/2} / \Gamma(d/2)$, where Γ is the Gamma function.

The density $f_{\mathbf{X}}$ can then be estimated by replacing g in (1) with (2), and denote the estimated $f_{\mathbf{X}}$ by $\widehat{f}_{n,h,a}$. The estimator defined in (2) depends on two tuning parameters: the bandwidth h , which is a classical tuning parameter for kernel smoothing, and the ‘‘extra’’ tuning parameter a , which helps to reduce the bias at 0. In the case $a = 0$, i.e., $\psi_a(\xi) = \xi$, this estimator reduces to the estimator introduced in [15].

Note that replacing $\boldsymbol{\Sigma}$ by $c\boldsymbol{\Sigma}$ and g by $g(c^{-1} \cdot)$ for some $c \in \mathbb{R}$ gives the same distribution, see, e.g., [9, Theorem 2.6.2]. To make the problem identifiable, the identifiability constraint $\boldsymbol{\Sigma} = \text{Var}[\mathbf{X}]$ is usually assumed. This has the implicit assumption that all components of \mathbf{X} have finite variance. Actually, we do not need such an assumption here, as we assume that $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are known, which then uniquely identifies the density generator g . Possibilities of extension of our procedures to the case where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are unknown are discussed in Sect. 6.1. In the rest of this chapter, we focus on the simpler case where no nuisance parameters ($\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$) are to be estimated, only the (functional) parameter g .

In order to use the estimator $\widehat{g}_{n,h,a}$, both tuning parameters h and a need to be chosen. Reference [11] shows that the estimator $\widehat{g}_{n,h,a}$ is uniformly convergent and asymptotically normal for any choice of a . However, the finite-distance behavior strongly depends on the choice of a , and also on the choice of h , as usual in kernel smoothing.

Therefore, in this chapter, we study the impact of the choice of h and a on the performance of $\widehat{g}_{n,h,a}$. In Sect. 2, we compute the MSE and discuss the optimal choice of h . Section 3 presents a new method for the estimation of the derivatives of the generator. New methods for choosing a are presented in Sect. 4. Finite-distance properties of the estimator are presented in a simulation study in Sect. 5.

The proposed methodology has been implemented and is available in the R package `ElliptCopulas` [6], in the functions `ElldistrEst.adapt` for the adaptive choice of a and h , and `ElldistrDerivEst` for the estimation of the derivatives.

Notations For a given generator g and a given $a > 0$, we define the functions \widetilde{g} , z_a , ρ_a by

$$\widetilde{g}(\xi) := \xi^{(d-2)/2} g(\xi), \quad z_a(\xi) := \widetilde{g}(\xi) / \psi'_a(\xi), \quad \text{and} \quad \rho_a := z_a \circ \psi_a^{-1}, \quad (3)$$

for any $\xi > 0$. These functions will be useful because the factor $\xi^{(d-2)/2}$ arises naturally when doing the change of variable to spherical coordinates (see Lemma 3 and the discussion therein). Note that ρ_a is the density of $\psi_a(\xi)$, see [11, p. 207].

2 MSE and Optimal Choice of h

We start by giving a first-order expansion for the both the bias and the variance of the estimator $\widehat{g}_{n,h,a}$. This theorem is proved in section “[Appendix 1: Proof of Theorem 2.1](#)”.

Theorem 2.1 *Let $h, a > 0$. Assume that ρ_a is three times continuously differentiable in a neighborhood of $\psi_a(\xi)$ for some $\xi > 0$. Then*

$$\text{Bias}_{n,h,a}(\xi) := \mathbb{E}[\widehat{g}_{n,h,a}(\xi)] - g(\xi) = h^2 \times \frac{\psi'_a(\xi) \rho''_a(\psi_a(\xi))}{2 \xi^{(d-2)/2}} \int_{\mathbb{R}} K(w) w^2 dw + O(h^3),$$

and

$$\text{Var}[\widehat{g}_{n,h,a}(\xi)] = \frac{\psi'_a(\xi) g(\xi) \|K\|_2^2}{n h s_d \xi^{(d-2)/2}} + o((nh)^{-1}),$$

as $h \rightarrow 0$ and $n \rightarrow \infty$.

We remark that the rates involved in the latter theorem do not depend on the dimension d . This confirms that such an elliptical assumption allows to avoid the curse of dimensionality. This theorem is similar to classical results on univariate kernel smoothing, featuring $\|K\|_2^2$ and $\int_{\mathbb{R}} K(w)w^2 dw$, as well as the rates h^2 for the bias and $1/(nh)$ for the variance, see, e.g., [16, Section 1.2].

The mean square error is then

$$\text{MSE}_{n,h,a}(\xi) := \mathbb{E} \left[\left(\widehat{g}_{n,h,a}(\xi) - g(\xi) \right)^2 \right] = \text{AMSE}_{n,h,a} + o(h^4) + o((nh)^{-1}),$$

where the asymptotic MSE is defined by

$$\text{AMSE}_{n,h,a} := \frac{C_1}{nh} \psi'_a(\xi) + C_2 \frac{h^4}{4} \left[\psi'_a(\xi) \cdot \rho''_a(\psi_a(\xi)) \right]^2,$$

for two positive constants C_1, C_2 defined by

$$C_1 := \frac{\|K\|_2^2}{s_d} \frac{g(\xi)}{\xi^{(d-2)/2}}, \quad C_2 := \frac{1}{\xi^{d-2}} \left(\int_{\mathbb{R}} K(w)w^2 dw \right)^2.$$

This asymptotic mean square error depends on the two tuning parameters h and a . Since this expression has a simpler dependence on h , we start by computing the optimal h that minimizes the AMSE for any given a .

Proposition 2.2 *Let $n > 0$, g be a generator, and $a > 0$. Let $A := a^{d/2}$. Then $h \mapsto \text{AMSE}_{n,h,a}$ is minimized for the choice h_{opt} defined by*

$$h_{\text{opt}}(n, a) := n^{-1/5} \left(\frac{C_1}{C_2} \right)^{1/5} \frac{\psi'_a(\xi)}{\left| \widetilde{\text{AMSE}}_A \right|^{2/5}}, \quad (4)$$

and

$$\text{AMSE}_{n,h_{\text{opt}},a} = n^{-4/5} \left((C_1 C_2)^{4/5} + C_1^{4/5} C_2^{1/5} \right) \left| \widetilde{\text{AMSE}}_A \right|^{2/5}, \quad (5)$$

where

$$\widetilde{\text{AMSE}}_A := \frac{\rho''_a(\psi_a(\xi))}{\psi'_a(\xi)} = \widetilde{g}''(\xi) + \frac{\widetilde{C}_3 A^2 + \widetilde{C}_4 A + \widetilde{C}_5}{(\xi^{d/2} + A)^2}, \quad (6)$$

and the constants are given by

$$\widetilde{C}_3 := \frac{d-2}{4\xi^2} (3(d-2)\widetilde{g}(\xi) - 6\xi\widetilde{g}'(\xi)),$$

$$\begin{aligned}\tilde{C}_4 &:= \frac{d-2}{4\xi^2} \left(\tilde{g}(\xi)(d-4) - 6\tilde{g}'(\xi)\xi^{(d+2)/2} \right), \\ \tilde{C}_5 &:= \frac{d^2-4}{4\xi^2} \tilde{g}(\xi).\end{aligned}$$

This proposition is proved in section “[Appendix 2: Proof of Proposition 2.2](#)”. Note that, for the optimal choice of h , the AMSE decouples the role of n and a . Indeed, Eq. (5) decomposes the AMSE with the bandwidth h_{opt} as a product of three terms: the classical rate of convergence $n^{-4/5}$, a constant term that only depends on $g(\xi)$, and the term $\widetilde{\text{AMSE}}_A$ that only depends on a and on the generator g . This means that any asymptotically optimal a_{opt} only depends on the normalized generator $\tilde{g}(\xi)$, and not on the sample size n , unlike the asymptotically optimal bandwidth h_{opt} . Furthermore, note that the asymptotically optimal bandwidth h_{opt} can also be decomposed into three terms, in the same way as the AMSE itself.

3 Estimation of the Derivatives of the Generator

Followed by Proposition 2.2, it appears that the AMSE depends on the derivatives of ρ . This means that we need to know the derivatives of the generator, since ρ depends on g . In practice, we will not know the function g and so its derivatives as well. Therefore, we decided to estimate the derivatives of g .

The construction of the estimators for the derivatives follows a similar method to [8] but adjusted for elliptical distributions. First, we fix two integers k and m such that $m \geq k + 2$ and $m - k$ is even. We introduce the estimator

$$\hat{\eta}_{k,n,h,a}(\xi) = \frac{1}{nh^{k+1}s_d} \sum_{i=1}^n \left[K_k \left(\frac{\psi_a(\xi) - \psi_a(\xi_i)}{h} \right) + K_k \left(\frac{\psi_a(\xi) + \psi_a(\xi_i)}{h} \right) \right],$$

where K_k is a kernel function of order (k, m) , i.e., K_k is assumed to satisfy

$$\int K_k(x)x^j dx = \begin{cases} 0 & \text{if } j = 0, 1, \dots, k-1, k+1, \dots, m-1, \\ (-1)^k k! & \text{if } j = k, \\ \mu_m(K_k) \neq 0 & \text{if } j = m. \end{cases}$$

As noted by [8], we can construct kernel K_k of order (k, m) by differentiating k times a kernel function K of order $m - k$.

Surprisingly, for $k > 0$ the estimator $\hat{\eta}_{k,n,h,a}(\xi)$ does not converge to a function of $g^{(k)}(\xi)$. Instead, it consistently estimates the unknown quantity $\rho_a^{(k)}(\psi_a(\xi))$. This claim is stated in the following proposition, and it is proved in section “[Appendix 5: Proof of Proposition 3.1: MSE for Estimating Derivatives](#)”.

Proposition 3.1 *Assume that the kernel K_k has a compact support and ρ_a is k -times continuously differentiable in a neighborhood of $\psi_a(\xi)$ for some $\xi > 0$. Then we have*

$$\mathbb{E}[\widehat{\eta}_{k,n,h,a}(\xi)] = \rho_a^{(k)}(\psi_a(\xi)) + \mu_m(K_k) \frac{h^{m-k}}{m!} \rho_a^{(m)}(\psi_a(\xi)) + O(h^{m-k+1}). \quad (7)$$

$$\text{Var}[\widehat{\eta}_{k,n,h,a}(\xi)] \sim \frac{g(\xi) \|K_k\|_2^2}{nh^{2k+1} s_d}. \quad (8)$$

However, it is possible to consistently estimate $g^{(k)}$ using $\widehat{\eta}_{0,n,h,a}, \dots, \widehat{\eta}_{k,n,h,a}$. This is a consequence of the following lemma.

Lemma 1 *For every $k \geq 0$, there exist $k + 1$ functions $\alpha_{0,k}(\xi), \dots, \alpha_{k,k}(\xi)$ such that*

$$\eta_{k,a}(\xi) := \rho_a^{(k)}(\psi_a(\xi)) = \sum_{i=0}^k \alpha_{i,k}(\xi) g^{(i)}(\xi),$$

and the functions $\alpha_{i,k}$ do not depend on g , but only on a and d . Furthermore, $\alpha_{k,k}(\xi) \neq 0$.

This lemma is proved in section “[Proof of Lemma 1](#)”, where a general formula for α is given in Eq. (33). By Lemma 1, we know that $[\eta_{0,a}, \dots, \eta_{k,a}]^\top = [\alpha_{i,j}]_{1 \leq i, j \leq k} \times [g, \dots, g^{(k)}]^\top$, where $\alpha_{i,j} := 0$ for $i > j$. In this sense, estimating the derivatives of g can be seen as a linear inverse problem with a triangular matrix. This allows for a straightforward estimation of $g^{(k)}$ by inverting the matrix $[\alpha_{i,j}]$. Indeed, one can write

$$g^{(k)} = \frac{1}{\alpha_{k,k}} \eta_{k,a} - \sum_{i=0}^{k-1} \frac{\alpha_{i,k}}{\alpha_{k,k}} g^{(i)},$$

from which a consistent estimator can be derived as

$$\widehat{g^{(k)}} := \frac{1}{\alpha_{k,k}} \widehat{\eta}_{k,n,h,a} - \sum_{i=0}^{k-1} \frac{\alpha_{i,k}}{\alpha_{k,k}} \widehat{g^{(i)}}.$$

As a particular case, for $k = 0$, $\rho_a^{(0)}(\psi_a(\xi)) = \xi^{(d-2)/2} g(\xi) / \psi_a'(\xi)$, hence $\alpha_{0,0}(\xi) = \xi^{(d-2)/2} / \psi_a'(\xi)$. We can recognize that the estimator $\widehat{g^{(0)}}$ is exactly Liebscher’s estimator, which was indeed introduced for the estimation of $g = g^{(0)}$. For $k = 1$, we obtain

$$\rho_a'(\psi_a(\xi)) = (\widetilde{g}(\xi) / \psi_a'(\xi))' \times \frac{1}{\psi_a'(\xi)}$$

$$= \frac{(d-2)\xi^{(d-4)/2}\psi'_a(\xi) - 2\xi^{(d-2)/2}\psi''_a(\xi)}{2\psi'_a(\xi)^3} \times g(\xi) + \frac{\xi^{(d-2)/2}}{\psi'_a(\xi)^2} \times g'(\xi).$$

This means that a consistent estimator of $g'(\xi)$ can be computed using the following estimator:

$$\begin{aligned} \widehat{g}'_{n,h,a}(\xi) &:= \left(\widehat{\eta}_{1,n,h,a}(\xi) - \frac{(d-2)\xi^{(d-4)/2}\psi'_a(\xi) - 2\xi^{(d-2)/2}\psi''_a(\xi)}{\psi'_a(\xi)^3} \times \widehat{g}_{n,h,a}(\xi) \right) \\ &\quad \times \frac{\psi'_a(\xi)^2}{\xi^{(d-2)/2}}. \end{aligned}$$

For $k = 2$, we can use the formula above, or equivalently, Lemma 6, which suggests to define the following consistent estimator of $g''(\xi)$:

$$\begin{aligned} \widehat{g}''_{n,h,a}(\xi) &:= \frac{\widehat{\eta}_{2,n,h,a}(\xi)[\psi'_a(\xi)]^3}{\xi^{(d-2)/2}} \\ &\quad + \frac{\psi''_a(\xi)[\xi\widehat{g}_{n,h,a}(\xi) + 2\xi\widehat{g}'_{n,h,a}(\xi) + (d-2)\widehat{g}_{n,h,a}(\xi)] + \widehat{g}_{n,h,a}(\xi)\psi'''_a(\xi)}{\xi\psi'_a(\xi)} \\ &\quad + \left(\frac{\psi''_a(\xi)}{\psi'_a(\xi)} \right)^2 \left(\widehat{g}'_{n,h,a}(\xi) - 2\widehat{g}_{n,h,a}(\xi) + \frac{d-2}{2} \frac{\widehat{g}_{n,h,a}(\xi)}{\xi} \right) \\ &\quad - (d-2) \frac{\widehat{g}'_{n,h,a}(\xi)}{\xi} - \frac{(d-2)(d-4)}{4} \frac{\widehat{g}_{n,h,a}(\xi)}{\xi^2}. \end{aligned}$$

4 Data-Driven Choice of the Tuning Parameters

4.1 Estimation of $\widetilde{\text{AMSE}}$

By Proposition 2.2, we know that $\widetilde{\text{AMSE}}_A = \rho''_a(\psi_a(\xi))/\psi'_a(\xi)$. Note that $\psi'_a(\xi)$ is a known function and that $\rho''_a(\psi_a(\xi))$ can be consistently estimated by $\widehat{\eta}_{k,n,h,a}(\xi)$. This suggests to introduce an estimator of $\widetilde{\text{AMSE}}$ defined by

$$\begin{aligned} \widehat{\widetilde{\text{AMSE}}}_{n,h,a} &:= \frac{\widehat{\eta}_{2,n,h,a}(\xi)}{\psi'_a(\xi)} \\ &= \frac{1}{nh_s d \psi'_a(\xi)} \sum_{i=1}^n \left[K_2 \left(\frac{\psi_a(\xi) - \psi_a(\xi_i)}{h} \right) + K_2 \left(\frac{\psi_a(\xi) + \psi_a(\xi_i)}{h} \right) \right], \end{aligned}$$

where K_2 is a kernel of order $(2, 2)$. As a consequence of Proposition 3.1, we obtain the asymptotic expansions for the bias and the variance of $\widehat{\text{AMSE}}_{n,h,a}$.

Corollary 4.1 *As $n \rightarrow \infty$, we have*

$$\mathbb{E}[\widehat{\text{AMSE}}] = \widetilde{\text{AMSE}}_A + \mu_m(K_2) \frac{h^2 \rho_a^{(4)}(\psi_a(\xi))}{4! \psi_a'(\xi)} + O(h^3),$$

and

$$\text{Var} \left[\widehat{\text{AMSE}} \right] \sim \frac{g(\xi) \|K_k\|_2^2}{nh^{2k+1} s_d (\psi_a'(\xi))^2}.$$

Remember that our goal is to find the value of a (or equivalently, of A) that minimizes $\widetilde{\text{AMSE}}$. For this, we fix a first-step bandwidth, denoted by h_1 , which gives a consistent estimator $\widetilde{\text{AMSE}}_{n,h_1,a}(\xi)$, that can be used as proxy for the true quantity $\widetilde{\text{AMSE}}$. Therefore, a value of a that minimizes the approximated criteria $\widehat{\text{AMSE}}$ will be close to the true minimizer a_{opt} of $\widetilde{\text{AMSE}}$. We denote by \widehat{a} the minimizer of $a \mapsto \widetilde{\text{AMSE}}_{n,h_1,a}(\xi)$ on a given grid a_1, \dots, a_N . From this approximate minimizer, we can then compute the (approximate) optimal bandwidth using Eq. (4). This equation requires the knowledge of $\widetilde{\text{AMSE}}$ (which we already estimated) and of C_1 .

C_1 can be estimated replacing $g(\xi)$ by the kernel-based estimator $\widehat{g}_{n,h_2,\widehat{a}}(\xi)$ for another first-step bandwidth h_2 . Note that we could use the same first-step bandwidth as for the estimation of $\widetilde{\text{AMSE}}$, i.e., $h_1 = h_2$. Nevertheless, this may not be a good idea in practice because the estimation of $\widetilde{\text{AMSE}}$ depends on the properties of $\widehat{\eta}_{2,n,h,a}$. On the contrary, the estimation of $g(\xi)$ depends on the properties of $\widehat{\eta}_{0,n,h,a}$ whose convergence rates (and therefore, optimal bandwidth) are different (see Proposition 3.1). From an estimator

$$\widehat{C}_1 := \frac{\|K\|_2^2 \widehat{g}_{n,h_2,\widehat{a}}}{s_d \xi^{(d-2)/2}} \quad (9)$$

of C_1 and $\widehat{\text{AMSE}}_{n,h_1,\widehat{a}}$, we can compute an approximate optimal bandwidth by

$$\widehat{h}(n, \widehat{a}) := n^{-1/5} \left(\frac{\widehat{C}_1}{C_2} \right)^{1/5} \frac{\psi_a'(\xi)}{\left| \widetilde{\text{AMSE}}_{n,h_1,\widehat{a}} \right|^{2/5}}. \quad (10)$$

Finally, we can compute Liebscher's estimator (2) of $g(\xi)$ using \widehat{h} and \widehat{a} . To sum up, we propose the following procedure detailed in Algorithm 1.

Algorithm 1: Choice of optimal tuning parameters and estimation of $g(\xi)$

Input: Dataset, first-step bandwidths h_1, h_2 , grid of possible values a_1, \dots, a_N
for $k \leftarrow 1$ **to** N **do**
 | Compute the approximate criteria $\widetilde{\text{AMSE}}_{n, h_1, a_k}$;
end
Compute $\widehat{a} := \arg \min_k |\widetilde{\text{AMSE}}_{n, h_1, a_k}|$;
Compute $\widehat{g}_{n, h_2, \widehat{a}}(\xi)$;
Compute \widehat{C}_1 from $\widehat{g}_{n, h_2, \widehat{a}}(\xi)$ by Eq. (9) ;
Compute $\widehat{h}(n, \widehat{a})$ by Eq. (10) ;
Output: An estimator $\widehat{g}_{n, \widehat{h}, \widehat{a}}(\xi)$ using the (approximate) optimal tuning parameters

4.2 Direct Optimization of $\widetilde{\text{AMSE}}$

Since $\partial A / \partial a = (d/2)a^{d/2-1} > 0$, we only need to investigate the critical point of $\widetilde{\text{AMSE}}_A$ to obtain a critical point of $\text{AMSE}(\widehat{g}_{n, h_{\text{opt}}, a}(\mathbf{x}))$. Note that it can also be minimized at $\widetilde{\text{AMSE}}_A = 0$. The following result is proved in section “[Appendix 3: Proof of Proposition 4.2](#)”.

Proposition 4.2 *Let $\Delta = \Delta(\xi) := \widetilde{g}(\xi) / \widetilde{g}'(\xi)$ and*

$$A^* = \frac{2(d+2)\Delta - \Delta(d-4)\xi^{(d-2)/2} + 6\xi^d}{6\xi^{d/2}(d-2)\Delta - \Delta(d-4) - 6\xi^{(d+2)/2}}. \quad (11)$$

There exists a computable value $A_{\text{opt}}(g) \in \{0, +\infty, A^\}$ depending only on $g(\xi)$, $g'(\xi)$ and d such that $A \mapsto \widetilde{\text{AMSE}}_A$ is minimized for A tending to $A_{\text{opt}}(g)$.*

Interestingly, the computation of the optimal tuning parameter a_{opt} (or equivalently, A_{opt}) only requires the knowledge of $g(\xi)$ and $g'(\xi)$, while the computation of the optimal value h_{opt} of the bandwidth also requires the knowledge of the second derivative $g''(\xi)$.

When we have chosen our value a , there remains to choose the tuning parameter h . For this, we can rely on the optimal h given by Eq. (4) which is

$$h_{\text{opt}}(n, a) := n^{-1/5} \left(\frac{C_1}{C_2} \right)^{1/5} \frac{\psi'_a(\xi)}{\widetilde{\text{AMSE}}_A^{2/5}}.$$

However, this still does depend on the knowledge of C_1 , which will not be available in practice. Since $C_1 := \frac{\Gamma(d/2)}{\pi^{d/2}} \xi^{(-d+2)/2} g(\xi) \|K\|_2^2$, it can be estimated by using \widehat{g} with a first-step bandwidth in order to get an estimate of $g(\xi)$. This problem does not exist for C_2 which only depends on the choice of the kernel K .

5 Simulation Study

In the previous sections, we have studied asymptotic properties of the estimators. As a complement to these theoretical results, we now empirically study the properties of the estimators for a finite sample size. For this, we use a Monte-Carlo approach with 200 replications and run these simulations on a few nodes on the DelftBlue supercomputer [3]. As a default family, we choose the Gaussian generator. To set a simple setting, we choose $\mu = 0$ and Σ to be the identity matrix. Studying the influence of the correlations is left for future research. It is expected that, outside of the boundary cases (correlations close to -1 or 1 or correlation matrix close to be not positive definite), the influence of the correlations should be limited. This was shown through simulation studies in a related setting for the estimation of elliptical copula density, see [4, Figures 8 and 9].

5.1 Influence of the Tuning Parameters on the MISE of Liebscher's Estimator $\hat{g}_{n,h,a}$ in Dimension 3

In the first simulation, we use a sample size $n = 5000$. On this sample, we apply Liebscher's estimator (2) with different values of the tuning parameters a and h . We study how these tuning parameters affect the mean integrated square error (MISE) defined by

$$\text{MISE} := \mathbb{E} \left[\int (\hat{g}(x) - g(x))^2 dx \right],$$

for a given estimator \hat{g} of g . We approximate the integral by a discrete sum over the points $\{i/10, i = 1, \dots, 50\}$. The corresponding results are displayed in Fig. 1. There seems to be an optimal combination of a and h that minimizes the MISE.

As expected, there exist values of h and a that minimize the MISE. Note that for each h there is a best value a_{opt} of a that minimizes the MISE. This optimal value of a depends on h and also on the sample size n (in general). To see the effect of the sample size n , we display the MISE as a function of n and h , where a is chosen as the optimal parameter. Figure 2 shows that the optimal tuning parameter h decreases as a function of n ; as expected the best MISE is also a decreasing function of the sample size.

To find more about the behavior of a_{opt} , we have displayed its value as a function of h and n in Fig. 3. We overlay the curve of $(n, h_{\text{opt}}(n))$ to give insight about which values of h are relevant. Coherently with our theoretical findings, the value of $a_{\text{opt}}(n, h_{\text{opt}}(n))$ seems roughly constant. This means that there is a "universal" value of the tuning parameter a which allows to reach the best MISE.

The optimal value of h is given by the balance between the bias term and the variance term (Fig. 4). This is the classical bias-variance trade-off in nonparametric

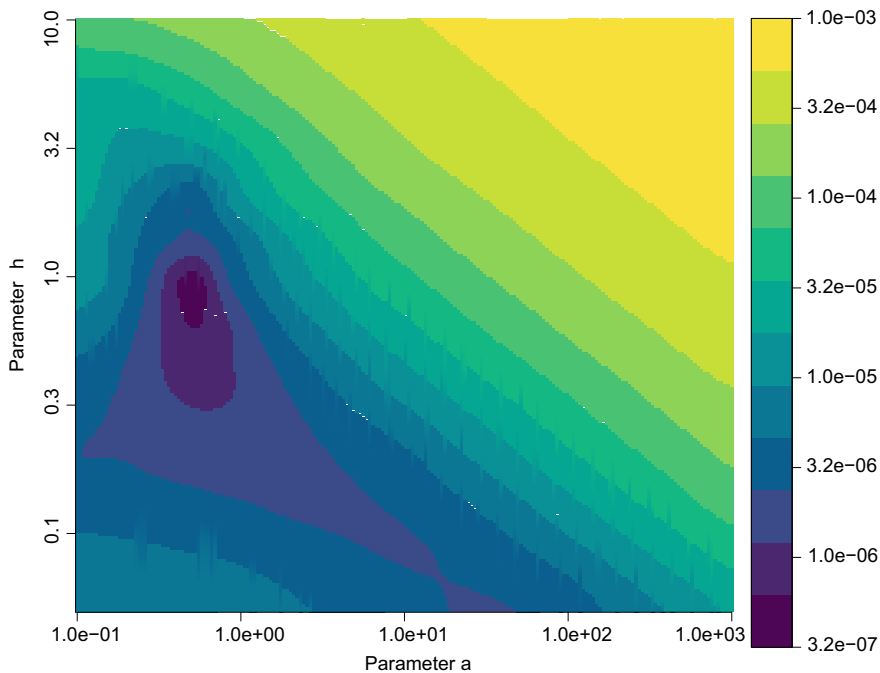


Fig. 1 MISE of Liebscher’s estimator $\hat{g}_{n,h,a}$ as a function of the tuning parameters a and h , for a sample size of 5000 using the Gaussian generator. All axes are in log scale; labels are given in the form $1.0e+01$ with the meaning 1.0×10^{01}

statistics [7]. Empirically, it seems that the best bandwidth in dimension 3 is $h(n) = 5.16 \times n^{-0.21}$. This is obtained by fitting a linear regression of $\log_{10}(n)$ on $\log_{10}(h(n))$, where $h(n)$ is the empirically found best bandwidth h for the sample size n . However, since we only considered the example of the Gaussian generator, this can only be considered a rule of thumb, and more research is needed to find under which conditions this rule of thumb is a good choice.

Similar phenomena can be observed for higher dimensions. We did the same simulations, but using dimensions $d \in \{4, 5, 8, 10\}$. The MISE as a function of the tuning parameters a and h is displayed for all four dimensions in Fig. 5. For all of them, a value of the parameter a around 1 is close to optimal.

5.2 Comparison of the First-Step Estimator with the New Estimator (Adaptive Choice of the Tuning Parameters)

In general, the optimal values of the tuning parameters h_{opt} and a_{opt} of Liebscher’s estimator depend on the true function g . In Algorithm 1, we propose a procedure

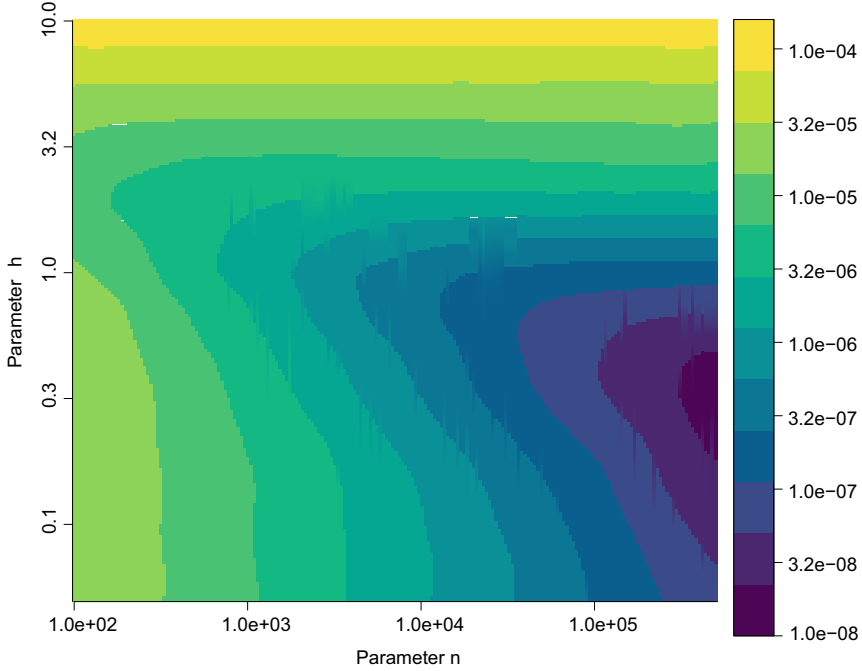


Fig. 2 MISE of Liebscher's estimator $\hat{g}_{n,h,a}$ as a function of the sample size n and the tuning parameter h . a is chosen as the best tuning parameter for each pair (n, h) . All axes are in log scale; labels are given in the form $1.0\text{e}+01$ with the meaning 1.0×10^{01} .

to determine the best tuning parameters. However, this procedure necessitates two first-step bandwidths h_1 and h_2 . In our simulation study and for simplicity, we choose these first-step bandwidths to be equal. We compare the MSE of the first-step estimator and the MSE of the final estimator for different choices of h and for different points x .

These results are displayed in Fig. 6. We can see that the performance of the final estimator is quite insensitive to the choice of h and allows to reach near-optimal performance without having to fine-tune the choice of h . This highlights an advantage of our methodology compared to the existing one. In our new method, the choice of the (first-step) bandwidth becomes less critical to obtain good performances.

The total computation of the simulations of the adaptive estimator was 693 hours. Depending on the sample size and the dimension, the computation time of the estimator given by Algorithm 1 can vary from 0.02 seconds to 70 seconds. Figure 7 illustrates the influence of the sample size and the dimension on the computation time, as expected.

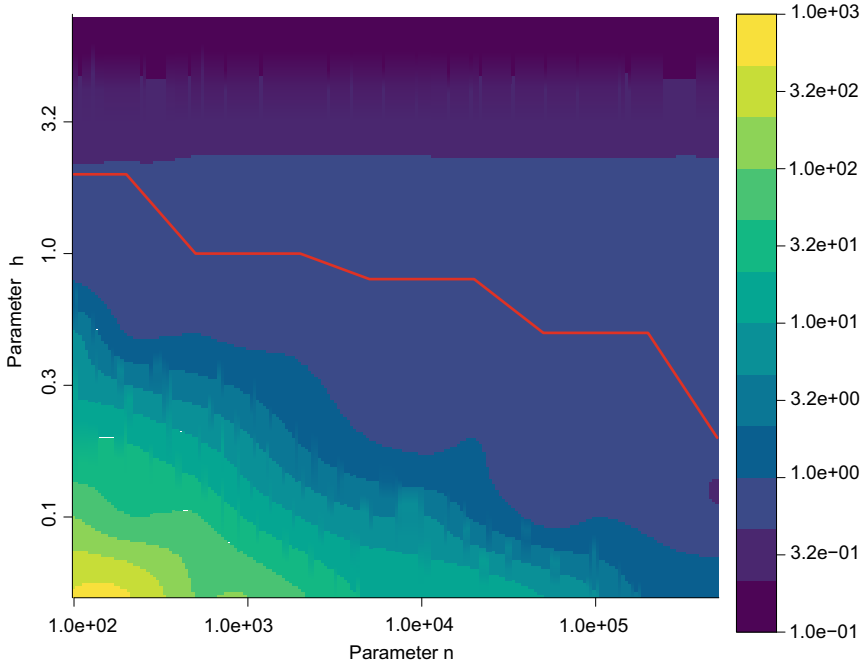


Fig. 3 Heatmap of the optimal tuning parameter a_{opt} as a function of the sample size n and the tuning parameter h . The red curve represents the best h as a function of n . All axes are in log scale; labels are given in the form $1.0e+01$ with the meaning 1.0×10^{01}

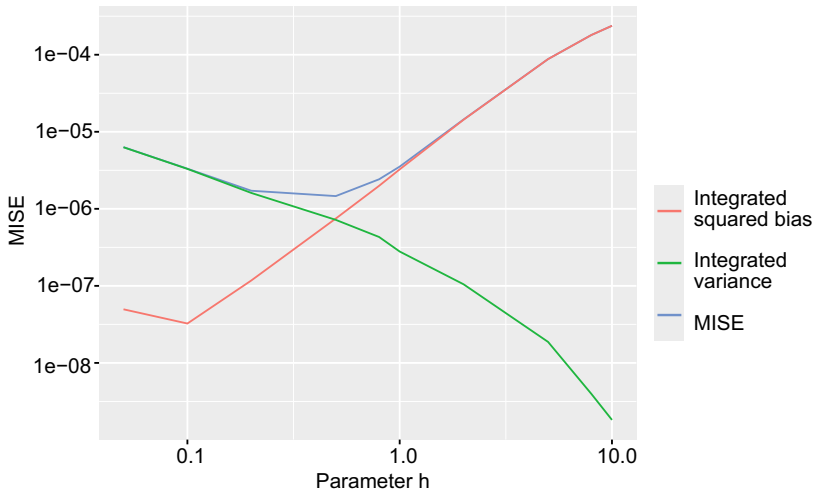


Fig. 4 Decomposition of the MISE of Liebscher’s estimator $\widehat{g}_{n,h,a}$ in terms of bias and variance, as a function of the tuning parameter h for $n = 5000$ and $a = 1$

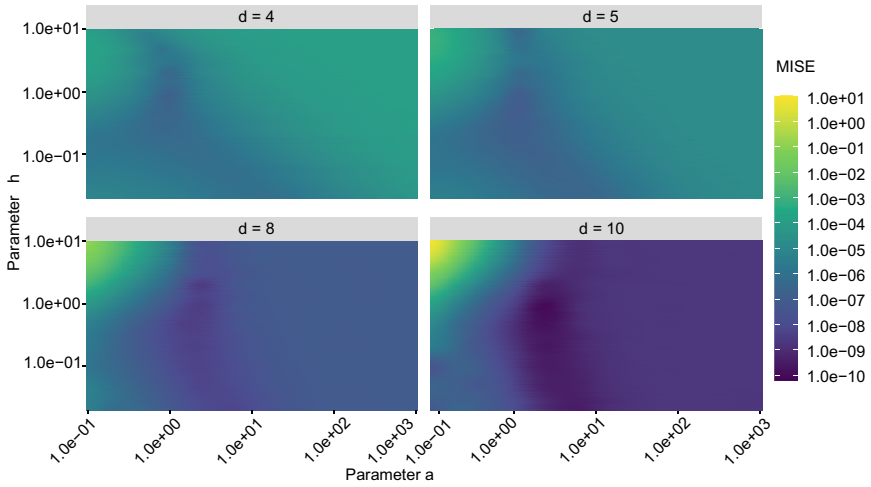


Fig. 5 MISE of Liebscher’s estimator $\hat{g}_{n,h,a}$ as a function of the tuning parameters a and h , for a sample size of 5000 using the Gaussian generator and for different dimensions

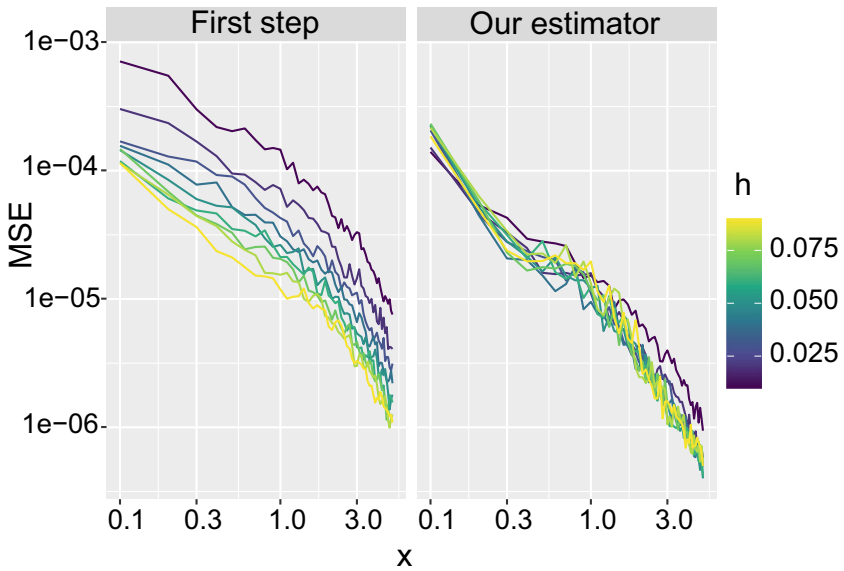


Fig. 6 Comparison of the MSE of the first-step estimator and of the estimator given by Algorithm 1 as a function of the tuning parameter h , for a sample size of 1000 using the Gaussian generator

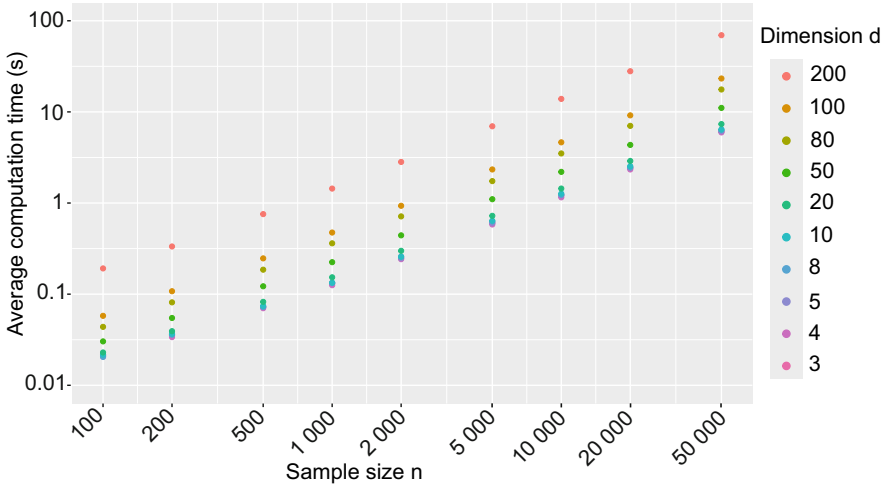


Fig. 7 Average computation time (in seconds) of the estimator given by Algorithm 1 as a function of the sample size n , for different dimensions

6 Extensions and Open Problems

6.1 Unknown μ and Σ

For technical convenience, we have assumed before that μ and Σ were known. In practice, this is not usually the case, and both need to be estimated. μ can be estimated via the sample mean (assuming that all the components of \mathbf{X} have a finite first moment), and Σ can be estimated via the sample covariance matrix (assuming that all the components of \mathbf{X} have a finite second moment). Robust estimators of μ and Σ are also available in the literature, see, e.g., [12].

If μ and Σ are replaced by estimators $\hat{\mu} = \hat{\mu}_{(\mathbf{x}_1, \dots, \mathbf{x}_n)}$ and $\hat{\Sigma} = \hat{\Sigma}_{(\mathbf{x}_1, \dots, \mathbf{x}_n)}$ in (2), we obtain a corresponding estimator $\hat{g}_{n,h,a}$ given by

$$\frac{\xi^{-\frac{d+2}{2}} \psi'_a(\xi)}{nhs_d} \sum_{i=1}^n K \left(\frac{\psi_a(\xi) - \psi_a \left(\frac{(\mathbf{x}_i - \hat{\mu}_{(\mathbf{x}_1, \dots, \mathbf{x}_n)})^\top \hat{\Sigma}_{(\mathbf{x}_1, \dots, \mathbf{x}_n)}^{-1} (\mathbf{x}_i - \hat{\mu}_{(\mathbf{x}_1, \dots, \mathbf{x}_n)})}{h} \right)}{h} \right) + K \left(\frac{\psi_a(\xi) + \psi_a \left(\frac{(\mathbf{x}_i - \hat{\mu}_{(\mathbf{x}_1, \dots, \mathbf{x}_n)})^\top \hat{\Sigma}_{(\mathbf{x}_1, \dots, \mathbf{x}_n)}^{-1} (\mathbf{x}_i - \hat{\mu}_{(\mathbf{x}_1, \dots, \mathbf{x}_n)})}{h} \right)}{h} \right),$$

whose bias necessitates the computation of integrals of the form

$$\int_{\mathbb{R}^{d(n-1)}} \int_{\mathbb{R}^d} K \left(\frac{\psi_a(\xi) - \psi_a \left((\mathbf{x}_i - \widehat{\boldsymbol{\mu}}_{(\mathbf{x}_1, \dots, \mathbf{x}_n)})^\top \widehat{\boldsymbol{\Sigma}}_{(\mathbf{x}_1, \dots, \mathbf{x}_n)}^{-1} (\mathbf{x}_i - \widehat{\boldsymbol{\mu}}_{(\mathbf{x}_1, \dots, \mathbf{x}_n)}) \right)}{h} \right) \\ \times |\boldsymbol{\Sigma}|^{-1/2} g \left((\mathbf{x}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) \right) d\mathbf{x}_i \prod_{\substack{1 \leq j \leq n \\ j \neq i}} f_{\mathbf{X}}(\mathbf{x}_j) d\mathbf{x}_j.$$

By the change of variables $\mathbf{z}_i := \widehat{\boldsymbol{\Sigma}}_{(\mathbf{x}_1, \dots, \mathbf{x}_n)}^{-1/2} (\mathbf{x}_i - \widehat{\boldsymbol{\mu}}_{(\mathbf{x}_1, \dots, \mathbf{x}_n)})$ and $\mathbf{z}_j = \mathbf{x}_j$ for $j \neq i$, the main factor in the integral can be rewritten as

$$K \left(\frac{\psi_a(\xi) - \psi_a(\|\mathbf{z}_i\|^2)}{h} \right) |\boldsymbol{\Sigma}|^{-1/2} g \left((\widehat{\boldsymbol{\Sigma}}^{1/2} \mathbf{z}_i + \widehat{\boldsymbol{\mu}} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} \right. \\ \left. \times (\widehat{\boldsymbol{\Sigma}}^{1/2} \mathbf{z}_i + \widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}) \right) d\mathbf{z}_i \times J,$$

where J is the Jacobian of the transformation. For usual estimators, this Jacobian could be computed, and if g is Lipschitz, this would lead to a bound on the bias. Such an analysis is quite complicated and left for future research.

Another possibility is to use a sample-splitting strategy (see, e.g., the discussion at the end of Section 6 of [5] which considers sample splitting for hypothesis testing). Here, we would use a certain proportion $p \in (0, 1)$ of the observations for the estimation of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ and the rest of the data for the estimation of g . This could make the theoretical analysis easier (conditioning on the first sample so that $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are given and then integrating with respect to them) and also introduces yet a third tuning parameter p to be chosen. This is also not easy. A reasonable choice for p is 50%; this was found to be nearly optimal in the simulation study done in [5, Section 6 of the supplement].

6.2 Estimation of g by Minimization of the Integrated MSE (MISE)

In general, minimizing the mean integrated squared error may not be possible because g itself need not to be square-integrable. To be an elliptical distribution generator, the only condition for g to satisfy is $\int t^{d/2-1} g(t) = 2/s_d$, where $s_d = 2\pi^{d/2}/\Gamma(d/2)$ (see, e.g., the introduction of [4]). Therefore, if $g(t) = C/(t\sqrt{t})\mathbf{1}_{t \leq 1}$ and $d = 4$, then the abovementioned condition can be satisfied since $t^{4/2-1}g(t) = C/\sqrt{t}$ is integrable on $[0, 1]$. On the contrary, $g^2(t) = 1/t^3$ is not integrable on $[0, 1]$. This means that

$$\text{MISE} = \int_0^{+\infty} (g - \widehat{g}_{n,h,a})^2 = +\infty,$$

since \widehat{g} is always square-integrable (as a finite sum of compactly supported bounded functions). Since the MISE is infinite for any choice of h , then this may not be the right tool for mathematical analysis.

Under some restrictions on g it may be possible to compute the MISE and to minimize it, but it would likely require the use of other mathematical techniques and is left for future research.

Acknowledgments The authors thank the two anonymous reviewers for their useful comments which significantly improved the manuscript.

Appendix 1: Proof of Theorem 2.1

Expression for the Bias

Step 1: Rewriting the Bias in a Simpler Form Because the random vectors $\mathbf{X}_1, \dots, \mathbf{X}_n$ are identically distributed, the expectation of $\widehat{g}_{n,h,a}$ can be written as

$$\mathbb{E}[\widehat{g}_{n,h,a}(\xi)] = \frac{\Gamma(d/2)}{\pi^{d/2}} \frac{\xi^{(-d+2)/2} \psi'_a(\xi)}{h} \left(\mathbb{E} \left[K \left(\frac{\psi_a(\xi) - \psi_a(\xi_1)}{h} \right) \right] + \mathbb{E} \left[K \left(\frac{\psi_a(\xi) + \psi_a(\xi_1)}{h} \right) \right] \right).$$

First, note that for h small enough,

$$\mathbb{E} \left[K \left(\frac{\psi_a(\xi) + \psi_a(\xi_1)}{h} \right) \right] = 0$$

by Lemma 2. Now, to compute the expectation of the other term, we apply Lemma 3, so that

$$\mathbb{E} \left[K \left(\frac{\psi_a(\xi) - \psi_a(\xi_1)}{h} \right) \right] = \frac{\pi^{d/2} h}{\Gamma(d/2)} \int_{-\psi_a(\xi)/h}^{\infty} K(w) \times \frac{g(\psi_a^{-1}(\psi_a(\xi) + wh)) \cdot \psi_a^{-1}(\psi_a(\xi) + wh)^{(d-2)/2}}{\psi'_a(\psi_a^{-1}(\psi_a(\xi) + wh))} dw.$$

By defining now the function

$$z(\xi) := \frac{g(\xi)}{\psi'_a(\xi)} \xi^{(d-2)/2}, \quad (12)$$

we get that

$$\begin{aligned} \text{Bias}_{h,a}(\xi) &= \xi^{(-d+2)/2} \psi'_a(\xi) \int_{-\psi_a(\xi)/h}^{\infty} K(w) z(\psi_a^{-1}(\psi_a(\xi) + wh)) dw - g(\xi) \\ &= \xi^{(-d+2)/2} \psi'_a(\xi) I_1 - g(\xi), \end{aligned}$$

where

$$\begin{aligned} I_1 &:= \int_{-\psi_a(\xi)/h}^{\infty} K(w) z(\psi_a^{-1}(\psi_a(\xi) + wh)) dw \\ &= \int_{-\infty}^{\infty} K(w) \cdot z(\psi_a^{-1}(\psi_a(\xi) + wh)) dw \\ &\quad - \int_{-\infty}^{-\psi_a(\xi)/h} K(w) \cdot z(\psi_a^{-1}(\psi_a(\xi) + wh)) dw. \end{aligned} \quad (13)$$

We have $-\psi_a(\xi)/h \rightarrow -\infty$ as $n \rightarrow +\infty$. Since we assume that K has compact support, we deduce that the second term in (13) is zero for n large enough by Lemma 2. Note that

$$g(\xi) = \frac{z(\xi) \psi'_a(\xi)}{\xi^{(d-2)/2}} = \frac{\psi'_a(\xi)}{\xi^{(d-2)/2}} \int_{-\infty}^{\infty} K(w) z(\xi) dw,$$

where the last equality comes from the assumption that $\int_{\mathbb{R}} K(w) dw = 1$. Therefore, we can rewrite the bias using the first term of (13), and replacing $g(\xi)$ as above,

$$\text{Bias}_{h,a}(\xi) = \xi^{(-d+2)/2} \psi'_a(\xi) \int_{-\infty}^{\infty} K(w) \left[z(\psi_a^{-1}(\psi_a(\xi) + wh)) - z(\xi) \right] dw.$$

Remark that $z(\psi_a^{-1}(\psi_a(\xi) + wh)) = (z \circ \psi_a^{-1})(\psi_a(\xi) + wh)$ and $z(\xi) = (z \circ \psi_a^{-1})(\psi_a(\xi))$. Therefore, since $\rho_a = z \circ \psi_a^{-1}$ by Eq. (3), we obtain

$$\text{Bias}_{h,a}(\xi) = \xi^{(-d+2)/2} \psi'_a(\xi) \int_{-\infty}^{\infty} K(w) [\rho_a(\psi_a(\xi) + wh) - \rho_a(\psi_a(\xi))] dw. \quad (14)$$

Step 2: Application of Taylor Expansion By applying the Taylor-Lagrange formula to the function ρ_a at the point $\psi_a(\xi)$, we get

$$\begin{aligned} \rho_a(\psi_a(\xi) + wh) - \rho_a(\psi_a(\xi)) &= \rho'_a(\psi_a(\xi)) wh + \frac{\rho''_a(\psi_a(\xi))}{2} (wh)^2 \\ &\quad + \frac{\rho'''_a(c_T)}{6} (wh)^3, \end{aligned}$$

where c_T lies between $\psi_a(\xi) + wh$ and $\psi_a(\xi)$. So, (14) becomes

$$\begin{aligned} \text{Bias}_{h,a}(\xi) &= \xi^{(-d+2)/2} \psi'_a(\xi) \int_{-\infty}^{\infty} K(w) \left[\rho'_a(\psi_a(\xi)) wh \right. \\ &\quad \left. + \frac{\rho''_a(\psi_a(\xi))}{2} (wh)^2 + \frac{\rho'''_a(c_T)}{2} (wh)^3 \right] dw \\ &= \frac{\psi'_a(\xi)}{\xi^{(d-2)/2}} \left(\rho'_a(\psi_a(\xi)) h \int_{\mathbb{R}} K(w) w dw \right. \\ &\quad \left. + h^2 \frac{\rho''_a(\psi_a(\xi))}{2} \int_{\mathbb{R}} K(w) w^2 dw + h^3 \int_{\mathbb{R}} K(w) \frac{\rho'''_a(c_T)}{6} w^3 dw \right). \end{aligned}$$

Note that $\int_{\mathbb{R}} K(w) w dw = 0$ since $K(w)$ is even and w is odd; their product is odd, and therefore the integral on \mathbb{R} is zero. As for

$$h^3 \int_{\mathbb{R}} K(w) \frac{\rho'''_a(c_T)}{6} w^3 dw,$$

the integral tends to a finite constant by Bochner's lemma. Indeed, using the assumption that ρ'''_a is continuous, we obtain that ρ'''_a is bounded on a neighborhood of $\psi_a(\|\mathbf{x}\|^2)$. We have now shown that

$$\text{Bias}_{h,a}(\xi) = h^2 \times \frac{\psi'_a(\xi) \rho''_a(\psi_a(\xi))}{2\xi^{(d-2)/2}} \int_{\mathbb{R}} K(w) w^2 dw + O(h^3).$$

Note that we can replace ρ''_a by its expression using z , and we get

$$\text{Bias}_{h,a}(\xi) = h^2 \times \frac{z''_a(\xi) \cdot \psi'_a(\xi) - \psi''_a(\xi) \cdot z'_a(\xi)}{2\xi^{(d-2)/2} [\psi'_a(\xi)]^2} \int_{\mathbb{R}} K(w) w^2 dw + O(h^3).$$

Expression for the Variance

We now study the variance term. By definition,

$$\text{Var}(\widehat{g}_{n,h,a}(\xi)) = \mathbb{E}[\widehat{g}_{n,h,a}(\xi)^2] - \mathbb{E}[\widehat{g}_{n,h,a}(\xi)]^2.$$

The term $\mathbb{E}[\widehat{g}_{n,h,a}(\xi)]^2$ is already known, so we focus on $\mathbb{E}[\widehat{g}_{n,h,a}(\mathbf{x})^2]$. Let

$$c(\xi) := \frac{\Gamma(d/2) \xi^{-d/2+1} \psi'_a(\xi)}{\pi^{d/2}}.$$

We have

$$\mathbb{E} \left[\widehat{g}_{n,h,a}(\xi)^2 \right] = \frac{c^2(\xi)}{n^2 h(n)^2} \mathbb{E} \left[\left(\sum_{i=1}^n \left[K \left(\frac{\psi_a(\xi) - \psi_a(\xi_i)}{h(n)} \right) + K \left(\frac{\psi_a(\xi) + \psi_a(\xi_i)}{h(n)} \right) \right] \right)^2 \right].$$

Expanding the sum, we find that

$$\mathbb{E} \left[\widehat{g}_{n,h,a}(\xi)^2 \right] = \frac{c^2(\xi)}{n^2 h(n)^2} \left(\sum_{i=1}^n (T_{1,i} + T_{2,i} + 2T_{3,i}) + \sum_{1 \leq i \neq j \leq n} (T_{4,i,j} + T_{5,i,j} + T_{6,i,j} + T_{7,i,j}) \right),$$

where

$$T_{1,i} := \mathbb{E} \left[K^2 \left(\frac{\psi_a(\xi) - \psi_a(\xi_i)}{h(n)} \right) \right] \quad (15)$$

$$T_{2,i} := \mathbb{E} \left[K^2 \left(\frac{\psi_a(\xi) + \psi_a(\xi_i)}{h(n)} \right) \right] \quad (16)$$

$$T_{3,i} := \mathbb{E} \left[K \left(\frac{\psi_a(\xi) - \psi_a(\xi_i)}{h(n)} \right) K \left(\frac{\psi_a(\xi) + \psi_a(\xi_i)}{h(n)} \right) \right] \quad (17)$$

$$T_{4,i,j} := \mathbb{E} \left[K \left(\frac{\psi_a(\xi) - \psi_a(\xi_i)}{h(n)} \right) K \left(\frac{\psi_a(\xi) - \psi_a(\xi_j)}{h(n)} \right) \right] \quad (18)$$

$$T_{5,i,j} := \mathbb{E} \left[K \left(\frac{\psi_a(\xi) - \psi_a(\xi_i)}{h(n)} \right) K \left(\frac{\psi_a(\xi) + \psi_a(\xi_j)}{h(n)} \right) \right] \quad (19)$$

$$T_{6,i,j} := \mathbb{E} \left[K \left(\frac{\psi_a(\xi) + \psi_a(\xi_i)}{h(n)} \right) K \left(\frac{\psi_a(\xi) - \psi_a(\xi_j)}{h(n)} \right) \right] \quad (20)$$

$$T_{7,i,j} := \mathbb{E} \left[K \left(\frac{\psi_a(\xi) + \psi_a(\xi_i)}{h(n)} \right) K \left(\frac{\psi_a(\xi) + \psi_a(\xi_j)}{h(n)} \right) \right]. \quad (21)$$

Therefore, because the random variables are identically distributed, we get

$$\mathbb{E} \left[\widehat{g}_{n,h,a}(\xi)^2 \right] = \frac{c^2(\xi)}{nh(n)^2} (T_{1,1} + T_{2,1} + 2T_{3,1} + (n-1)T_{4,1,2} + (n-1)T_{5,1,2} + (n-1)T_{6,1,2} + (n-1)T_{7,1,2}).$$

Step 1. Removing Terms That Disappear We start by applying Lemma 2. This means that for all n large enough,

$$\mathbb{E} \left[\widehat{g}_{n,h,a}(\xi)^2 \right] = \frac{c^2(\xi)}{nh(n)^2} (T_{1,1} + (n-1)T_{4,1,2}).$$

Step 2. Computation of an Equivalent of $T_{1,1}$ We apply Lemma 3, with K^2 instead of K , and obtain

$$\begin{aligned} T_{1,1} &= \frac{\pi^{d/2}h(n)}{\Gamma(d/2)} \\ &\times \int_{-\psi_a(\xi)/h}^{\infty} K^2(w) \frac{g(\psi_a^{-1}(\psi_a(\xi) + wh(n))) \cdot \psi_a^{-1}(\psi_a(\xi) + wh)^{(d-2)/2}}{\psi'_a(\psi_a^{-1}(\psi_a(\xi) + wh(n)))} dw \end{aligned}$$

and so

$$\begin{aligned} \frac{c^2(\xi)}{nh(n)^2} T_{1,1} &= \frac{1}{nh(n)} \frac{\Gamma(d/2)\xi^{-d+2}\psi'_a(\xi)^2}{\pi^{d/2}} \\ &\times \int_{w(0)}^{\infty} K^2(w) \frac{g(\psi_a^{-1}(\psi_a(\xi) + wh(n))) \cdot \psi_a^{-1}(\psi_a(\xi) + wh)^{(d-2)/2}}{\psi'_a(\psi_a^{-1}(\psi_a(\xi) + wh(n)))} dw \\ &\sim \frac{1}{nh(n)} \frac{\Gamma(d/2)}{\pi^{d/2}} \xi^{(-d+2)/2} \psi'_a(\xi) g(\xi) \int_{-\infty}^{\infty} K^2(w) dw, \end{aligned}$$

by Bochner's lemma.

Step 3. Computation of $T_{4,1,2}$ Since the $X_i, i = 1, \dots, n$, are independent and identically distributed, we have

$$\begin{aligned} T_{4,1,2} &= \mathbb{E} \left[K \left(\frac{\psi_a(\xi) - \psi_a(\xi_1)}{h(n)} \right) \right] \cdot \mathbb{E} \left[K \left(\frac{\psi_a(\xi) - \psi_a(\xi_2)}{h(n)} \right) \right] \\ &= \mathbb{E} \left[K \left(\frac{\psi_a(\xi) - \psi_a(\xi_1)}{h(n)} \right) \right]^2 \\ &= \mathbb{E} \left[K \left(\frac{\psi_a(\xi) - \psi_a(\xi_1)}{h(n)} \right) \right]^2 \\ &= \frac{h(n)^2}{c^2(\xi)} \mathbb{E} \left[\widehat{g}_{n,h,a}(\xi) \right]^2, \end{aligned}$$

for n large enough by Lemma 2.

Step 4. Combining All Previous Results

$$\begin{aligned}
\text{Var}(\widehat{g}_{n,h,a}(\xi)) &= \mathbb{E}[\widehat{g}_{n,h,a}(\xi)^2] - \mathbb{E}[\widehat{g}_{n,h,a}(\xi)]^2 \\
&= \frac{c^2(\xi)}{nh(n)^2} (T_{1,1} + (n-1)T_{4,1,2}) - \mathbb{E}[\widehat{g}_{n,h,a}(\xi)]^2 \\
&= \frac{c^2(\xi)}{nh(n)^2} (T_{1,1} + (n-1)T_{4,1,2}) - (g(\xi) - \text{Bias}_{h,a}(\xi))^2 \\
&= \frac{n-1}{n} (g(\xi) - \text{Bias}_{h,a}(\xi))^2 + \frac{c^2(\xi)}{nh(n)^2} T_{1,1} \\
&\quad - (g(\xi) - \text{Bias}_{h,a}(\xi))^2 \\
&= -\frac{1}{n} (g(\xi) - \text{Bias}_{h,a}(\xi))^2 + \frac{c^2(\xi)}{nh(n)^2} T_{1,1}.
\end{aligned}$$

Since $\text{Bias}_{h,a}(\xi) = o(1)$, we get

$$\begin{aligned}
\text{Var}(\widehat{g}_{n,h,a}(\xi)) &= \frac{c^2(\xi)}{nh(n)^2} T_{1,1} + O(1/n) \\
&= \frac{1}{nh(n)} \frac{\Gamma(d/2)}{\pi^{d/2}} \xi^{(-d+2)/2} \psi'_a(\xi) g(\xi) \|K\|_2^2 \\
&\quad + o(1/(nh(n))) + O(1/n)
\end{aligned}$$

since $h(n) \rightarrow 0$ as $n \rightarrow \infty$.

Appendix 2: Proof of Proposition 2.2

We start from the expression of the asymptotic mean square error

$$\text{AMSE}_{n,h,a} := \frac{C_1}{nh} \psi'_a(\xi) + C_2 \frac{h^4}{4} \psi'_a(\xi)^2 \rho_a''^2(\psi_a(\xi)).$$

The partial derivative of $\text{AMSE}_{n,h,a}$ with respect to h is

$$\frac{\partial}{\partial h} \text{AMSE}_{n,h,a} = -\frac{C_1}{nh^2} \psi'_a(\xi) + C_2 h^3 \psi'_a(\xi)^2 \rho_a''^2(\psi_a(\xi)).$$

Solving $\frac{\partial}{\partial h} \text{AMSE}_{n,h,a} = 0$ for h gives the optimal bandwidth

$$h_{\text{opt}}(n) = \varphi_a(\xi) n^{-1/5}, \quad (22)$$

where

$$\varphi_a(\xi) := \left(\frac{C_1}{C_2 \psi'_a(\xi) \rho_a''(\psi_a(\xi))} \right)^{1/5} = C_3 \frac{\psi'_a}{(z'_a \psi'_a - \psi''_a z'_a)^{2/5}}, \quad (23)$$

and $C_3 := (C_1/C_2)^{1/5}$. The second equality in Eq. (23) is a consequence of Lemma 4, where the function z_a is defined by $z_a(\xi) := \frac{g(\xi)}{\psi'_a(\xi)} \xi^{(d-2)/2}$. Substituting h_{opt} by its expression (22) in the AMSE gives

$$\begin{aligned} \text{AMSE}_{n, h_{\text{opt}}, a} &= \frac{C_1}{\varphi_a(\xi) n^{4/5}} \psi'_a(\xi) + C_2 \varphi_a(\xi)^4 n^{-4/5} \psi'_a(\xi)^2 \rho_a''(\psi_a(\xi)) \\ &= n^{-4/5} \left(C_1 \frac{\psi'_a}{\varphi_a} + C_2 \varphi_a^4 \psi_a''(\rho_a'' \circ \psi_a) \right). \end{aligned}$$

By Eq. (23), we get that $\psi'_a/\varphi_a = C_3^{-1} (z'_a \psi'_a - \psi''_a z'_a)^{2/5}$. By Lemma 4, we also obtain

$$\begin{aligned} \varphi_a^4 \psi_a''(\rho_a'' \circ \psi_a) &= C_3^4 \frac{\psi_a'^4}{(z'_a \psi'_a - \psi''_a z'_a)^{8/5}} \psi_a'' \left(\frac{z'_a \psi'_a - \psi''_a z'_a}{\psi_a'^3} \right)^2 \\ &= C_3^4 (z'_a \psi'_a - \psi''_a z'_a)^{2/5}. \end{aligned}$$

Therefore,

$$\begin{aligned} \text{AMSE}_{n, h_{\text{opt}}, a} &= n^{-4/5} \left(\frac{C_1}{C_3} + C_2 C_3^4 \right) \frac{\rho_a''(\psi_a(\xi))^{2/5}}{\psi_a'(\xi)^{2/5}} \\ &= n^{-4/5} \left(\frac{C_1}{C_3} + C_2 C_3^4 \right) (z'_a \psi'_a - \psi''_a z'_a)^{2/5} \\ &= n^{-4/5} \left((C_1 C_2)^{4/5} + C_1^{4/5} C_2^{1/5} \right) (z'_a \psi'_a - \psi''_a z'_a)^{2/5} \\ &= n^{-4/5} \left((C_1 C_2)^{4/5} + C_1^{4/5} C_2^{1/5} \right) \\ &\quad \times \left(\tilde{g}''(\xi) + \frac{\tilde{C}_3 A^2 + \tilde{C}_4 A + \tilde{C}_5}{(\xi^{d/2} + A)^2} \right)^{2/5}, \end{aligned}$$

by Lemma 7. The expression of h_{opt} is obtained by combining Eqs. (22) and (23) and Lemma 7.

Appendix 3: Proof of Proposition 4.2

In Proposition 2.2, we show that the asymptotic MSE with optimal bandwidth h has the following expression:

$$\begin{aligned} \text{AMSE}_{n,h_{\text{opt}},a} &= n^{-4/5} \left((C_1 C_2)^{4/5} + C_1^{4/5} C_2^{1/5} \right) \\ &\quad \times \left(\tilde{g}''(\|\mathbf{x}\|^2) + \frac{\tilde{C}_3 A^2 + \tilde{C}_4 A + \tilde{C}_5}{(\xi^{d/2} + A)^2} \right)^{2/5}, \end{aligned}$$

where $A = a^{d/2}$. Therefore, we can minimize the function $a \mapsto \text{AMSE}_{n,h_{\text{opt}},a}$ by finding

$$A_{\text{opt}} := \operatorname{argmin}_{A>0} \left\{ \frac{\tilde{C}_3 A^2 + \tilde{C}_4 A + \tilde{C}_5}{(\xi^{d/2} + A)^2} \right\},$$

and so $a_{\text{opt}} := A_{\text{opt}}^{2/d}$ minimizes $a \mapsto \text{AMSE}_{n,h_{\text{opt}},a}$. In this section, we analyze in which cases the minimum exists.

Followed from the definition of A_{opt} , we define the following function:

$$\kappa(A) := \frac{\tilde{C}_3 A^2 + \tilde{C}_4 A + \tilde{C}_5}{(\xi^{d/2} + A)^2}. \quad (24)$$

Note that κ is a nonconstant ratio of polynomials, and therefore it is an analytic function. Therefore, there exists a nonzero higher order derivative of κ . For the first derivative of κ , we have

$$\begin{aligned} \kappa'(A) &= \frac{(2\tilde{C}_3 A + \tilde{C}_4)(\xi^{d/2} + A)^2 - (2\xi^{d/2} + 2A)(\tilde{C}_3 A^2 + \tilde{C}_4 A + \tilde{C}_5)}{(\xi^{d/2} + A)^4} \\ &= \frac{(2\tilde{C}_3 A + \tilde{C}_4)(\xi^{d/2} + A) - 2(\tilde{C}_3 A^2 + \tilde{C}_4 A + \tilde{C}_5)}{(\xi^{d/2} + A)^3} \\ &= \frac{\tilde{C}_4 \xi^{d/2} + \tilde{C}_4 A + 2\tilde{C}_3 \xi^{d/2} A - 2\tilde{C}_4 A - 2\tilde{C}_5}{(\xi^{d/2} + A)^3} \\ &= \frac{A(2\tilde{C}_3 \xi^{d/2} - \tilde{C}_4) + \tilde{C}_4 \xi^{d/2} - 2\tilde{C}_5}{(\xi^{d/2} + A)^3}. \end{aligned}$$

We again differentiate κ w.r.t A to get

$$\kappa''(A) = \frac{(2\tilde{C}_3 \xi^{d/2} - \tilde{C}_4)(\xi^{d/2} + A)^3 - 3(\xi^{d/2} + A)^2 (A(2\tilde{C}_3 \xi^{d/2} - \tilde{C}_4) + \tilde{C}_4 \xi^{d/2} - 2\tilde{C}_5)}{(\xi^{d/2} + A)^6}$$

$$\begin{aligned}
&= \frac{(2\tilde{C}_3\xi^{d/2} - \tilde{C}_4)(\xi^{d/2} + A) - 3(A(2\tilde{C}_3\xi^{d/2} - \tilde{C}_4) + \tilde{C}_4\xi^{d/2} - 2\tilde{C}_5)}{(\xi^{d/2} + A)^4} \\
&= \frac{-2A(2\tilde{C}_3\xi^{d/2} - \tilde{C}_4) + 2\tilde{C}_3\xi^d - 4\tilde{C}_4\xi^{d/2} + 6\tilde{C}_5}{(\xi^{d/2} + A)^4}.
\end{aligned}$$

From κ' , we determine for which A is κ maximum or minimum. We have the following cases:

Case 1. If $2\tilde{C}_3\xi^{d/2} - \tilde{C}_4 = 0$, then we consider the following three subcases.

Subcase 1a. If $\tilde{C}_4\xi^{d/2} - 2\tilde{C}_5 < 0$, then $\kappa' < 0$. The function κ decreases, and its minimum is attained when $A \rightarrow +\infty$.

Subcase 1b. If $\tilde{C}_4\xi^{d/2} - 2\tilde{C}_5 = 0$, then $\kappa' \equiv 0$. This means that κ is a constant function.

Subcase 1c. If $\tilde{C}_4\xi^{d/2} - 2\tilde{C}_5 > 0$, then $\kappa > 0$. Hence κ reaches its minimum as $A \downarrow 0$, since κ is an increasing function in this subcase.

Case 2. Suppose $2\tilde{C}_3\xi^{d/2} - \tilde{C}_4 \neq 0$. Let $A^* = (2\tilde{C}_5 - \tilde{C}_4\xi^{d/2})/(2\tilde{C}_3\xi^{d/2} - \tilde{C}_4)$; then $\kappa'(A^*) = 0$. Therefore,

$$\begin{aligned}
\kappa''(A^*) &= \frac{-2(2\tilde{C}_5 - \tilde{C}_4\xi^{d/2}) + 2\tilde{C}_3\xi^d - 4\tilde{C}_4\xi^{d/2} + 6\tilde{C}_5}{(C_6 + A^*)^4} \\
&= \frac{2\tilde{C}_5 + 2\tilde{C}_3\xi^d - 6\tilde{C}_4\xi^{d/2}}{(\xi^{d/2} + A^*)^4}.
\end{aligned}$$

Now we determine whether the (local) extrema of κ is attained at A^* , and we also have three subcases to consider.

Subcase 2a. If $2\tilde{C}_5 + 2\tilde{C}_3\xi^d - 6\tilde{C}_4\xi^{d/2} > 0$, then $\kappa''(A^*) > 0$. There is only one extremum, so the function κ has a global minimum at A^* .

Subcase 2b. If $2\tilde{C}_5 + 2\tilde{C}_3\xi^d - 6\tilde{C}_4\xi^{d/2} < 0$, then $\kappa''(A^*) < 0$. This means that A^* is a local maximum. However, there is only one extremum and so A^* is a global maximum. The minimum is attained at either as $A \downarrow 0$ or as $A \rightarrow \infty$.

Subcase 2c. If $2\tilde{C}_5 + 2\tilde{C}_3\xi^d - 6\tilde{C}_4\xi^{d/2} = 0$, then the minimum could be attained at 0 , $+\infty$, or A^* .

The expressions for $2\tilde{C}_3\xi^{d/2} - \tilde{C}_4$, $\tilde{C}_4\xi^{d/2} - 2\tilde{C}_5$, and $2\tilde{C}_5 + 2\tilde{C}_3\xi^d - 6\tilde{C}_4\xi^{d/2}$ are in Lemma 8.

From these cases, we conclude that $a \mapsto \text{AMSE}_{n, h_{\text{opt}}, a}$ is minimized when $a \downarrow 0$, $a \rightarrow \infty$, or at $a = (A^*)^{2/d}$. The only case where we can compute a_{opt} explicitly is in Subcase 2a. Recall that

$$A^* = \frac{2\tilde{C}_5 - \tilde{C}_4\xi^{d/2}}{2\tilde{C}_3\xi^{d/2} - \tilde{C}_4}.$$

Using the definitions of the constants given in Proposition 2.2, we have

$$\begin{aligned}
A^* &= \frac{2 \left[\frac{d-2}{4\xi^2} (d+2) \tilde{g}(\xi) \right] - \frac{d-2}{4\xi^2} \left[\tilde{g}(\xi)(d-4) - 6 \tilde{g}'(\xi) \xi^{(d+2)/2} \right] \xi^{d/2}}{2 \left[\frac{d-2}{4\xi^2} (3(d-2) \tilde{g}(\xi) - 6 \tilde{g}'(\xi) \xi) \right] \xi^{d/2} - \frac{d-2}{4\xi^2} \left[\tilde{g}(\xi)(d-4) - 6 \tilde{g}'(\xi) \xi^{(d+2)/2} \right]} \\
&= \frac{2 \left[(d+2) \tilde{g}(\xi) \right] - \left[\tilde{g}(\xi)(d-4) - 6 \tilde{g}'(\xi) \xi^{(d+2)/2} \right] \xi^{d/2}}{2 \left[(3(d-2) \tilde{g}(\xi) - 6 \tilde{g}'(\xi) \xi) \right] \xi^{d/2} - \left[\tilde{g}(\xi)(d-4) - 6 \tilde{g}'(\xi) \xi^{(d+2)/2} \right]} \\
&= \frac{\left[2(d+2) - (d-4) \xi^{d/2} \right] \tilde{g}(\xi) + 6 \xi^{d+1} \tilde{g}'(\xi)}{\left[6(d-2) \xi^{d/2} - d + 4 \right] \tilde{g}(\xi) - 6 \xi^{(d+2)/2} \tilde{g}'(\xi)}.
\end{aligned}$$

Let $\Delta(\xi) := \tilde{g}(\xi) / \tilde{g}'(\xi)$; then

$$A^* = \frac{\Delta(\xi) \left[2d + 4 - (d-4) \xi^{d/2} \right] + 6 \xi^{d+1}}{\Delta(\xi) \left[6 \xi^{d/2} (d-2) - d + 4 \right] - 6 \xi^{(d+2)/2}},$$

which finishes the proof of Proposition 4.2.

Appendix 4: Technical Results

Computation of Integrals

Lemma 2 *Let ψ be a measurable function from \mathbb{R}^+ to \mathbb{R}^+ and $\xi > 0$ such that $\psi(\xi) > 0$. Assume that the kernel K has compact support and $h(n) \rightarrow 0$ as $n \rightarrow \infty$. Let Z be a bounded random variable. Let I_n be defined by*

$$I_n := \mathbb{E} \left[Z \cdot K \left(\frac{\psi(\xi) + \psi((\mathbf{X}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{X}_i - \boldsymbol{\mu}))}{h(n)} \right) \right].$$

Then there exists $N \in \mathbb{N}$ such that $I_n = 0$ for all $n \geq N$.

Proof Let $[a, b]$ be the support of K , and remember $\xi_i := (\mathbf{X}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{X}_i - \boldsymbol{\mu})$. We have

$$I_n := \mathbb{E} \left[Z \cdot K \left(\frac{\psi(\xi) + \psi(\xi_i)}{h(n)} \right) \right] = \int z \cdot K \left(\frac{\psi(\xi) + \psi(x)}{h(n)} \right) d\mathbb{P}(x, z),$$

where \mathbb{P} is the law of the random vector (ξ_i, Z) . For any $x > 0$,

$$\frac{\psi(\xi) + \psi(x)}{h(n)} \geq \frac{\psi(\xi)}{h(n)} \rightarrow +\infty.$$

Therefore, for all n large enough, we have

$$\text{for all } x > 0, \frac{\psi(\xi) + \psi(x)}{h(n)} > b,$$

and therefore

$$\text{for all } x > 0, K\left(\frac{\psi(\xi) + \psi(x)}{h(n)}\right) = 0.$$

Finally, we get $I_n = 0$ for n large enough, as claimed. □

In many computations involving elliptical distributions, one can use a well-chosen change of variable to transform an expectation over the elliptically distributed variable \mathbf{X}_1 (i.e., a d -dimensional integral) into a one-dimensional integral. This is formalized in the following lemma, which is a consequence of the results given in [15]. They show that by using d -dimensional spherical coordinates, we can simplify these integrals (see pages 174–175 and Lemma 4.3 of [15]).

Lemma 3 *For any $\xi > 0$, for any $a, h > 0$, and for any function $K : \mathbb{R} \rightarrow \mathbb{R}$, we have*

$$\begin{aligned} \mathbb{E}\left[K\left(\frac{\psi_a(\xi) - \psi_a(\xi_1)}{h}\right)\right] &= \frac{\pi^{d/2}h}{\Gamma(d/2)} \int_{-\psi_a(\xi)/h}^{\infty} K(-w) \times \\ &\quad \frac{g(\psi_a^{-1}(\psi_a(\|\mathbf{x}\|^2) + wh)) \cdot \psi_a^{-1}(\psi_a(\xi) + wh)^{(d-2)/2}}{\psi_a'(\psi_a^{-1}(\psi_a(\|\mathbf{x}\|^2) + wh))} dw. \end{aligned}$$

Proof We begin by doing the change of variables $\mathbf{z}_i := \Sigma^{-1/2}(\mathbf{x}_i - \boldsymbol{\mu})$, and then we apply Stute and Werner’s [15] change of variable using the d -dimensional spherical coordinates, so that

$$\begin{aligned} &\mathbb{E}\left[K\left(\frac{\psi_a(\xi) - \psi_a(\xi_1)}{h}\right)\right] \\ &= \mathbb{E}\left[K\left(\frac{\psi_a(\xi) - \psi_a((\mathbf{X}_i - \boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{X}_i - \boldsymbol{\mu}))}{h}\right)\right] \\ &= \int_{\mathbb{R}^d} K\left(\frac{\psi_a(\xi) - \psi_a(\|\mathbf{z}_i\|^2)}{h}\right) |\Sigma|^{-1/2} g(\|\mathbf{z}_i\|^2) |\Sigma|^{1/2} d\mathbf{z}_i \\ &= \frac{\pi^{d/2}}{\Gamma(d/2)} \int_0^\infty K\left(\frac{\psi_a(\xi) - \psi_a(v_1)}{h}\right) g(v_1) v_1^{(d-2)/2} dv_1. \end{aligned}$$

We now do a third change of variables

$$w := w(v_1) = \frac{-\psi_a(\xi) + \psi_a(v_1)}{h} \iff v_1(w) = \psi_a^{-1}(\psi_a(\xi) + wh),$$

which gives

$$\mathbb{E} \left[K \left(\frac{\psi_a(\xi) - \psi_a(\xi_i)}{h} \right) \right] = \frac{\pi^{d/2} h}{\Gamma(d/2)} \int_{w(0)}^{\infty} K(-w) \times \frac{g(\psi_a^{-1}(\psi_a(\xi) + wh)) \cdot v_1(w)^{(d-2)/2}}{\psi'_a(\psi_a^{-1}(\psi_a(\xi) + wh))} dw.$$

Noting that $w(0) = -\psi_a(\xi)/h$ finishes the proof. \square

Technical Computations

We remind that z_a is the function defined by $z_a(\xi) := \frac{g(\xi)}{\psi'_a(\xi)} \xi^{(d-2)/2}$.

Lemma 4 *For any integer $d > 0$ and any $\xi, a > 0$, we have*

$$\rho''_a(\psi_a(\xi)) = \frac{z''_a(\xi) \cdot \psi'_a(\xi) - \psi''_a(\xi) \cdot z'_a(\xi)}{[\psi'_a(\xi)]^3}.$$

Proof Let $\Psi_a(t)$ be the inverse function of ψ_a , and let Ψ'_a be the derivative of Ψ_a w.r.t. t . Recall that $\rho_a = z_a \circ \Psi_a$. By differentiating this expression, we obtain $\rho'_a = (z'_a \circ \Psi_a) \cdot \Psi'_a$ and

$$\rho''_a = (z'_a \circ \Psi_a)' \cdot \Psi'_a + (z'_a \circ \Psi_a) \cdot \Psi''_a = (z''_a \circ \Psi_a) \cdot (\Psi'_a)^2 + (z'_a \circ \Psi_a) \cdot \Psi''_a.$$

So,

$$\rho''_a(\psi_a(\xi)) = z''_a(\xi) \cdot (\Psi'_a)^2(\psi_a(\xi)) + z'_a(\xi) \cdot \Psi''_a(\psi_a(\xi)). \quad (25)$$

Since

$$\Psi'_a(t) = \frac{1}{\psi'(\Psi_a(t))} \quad \text{and} \quad \Psi''_a(t) = -\frac{\psi''(\Psi_a(t))}{(\psi'_a(\Psi_a(t)))^3},$$

we get

$$(\Psi'_a)^2(\psi_a(\xi)) = \frac{1}{[\psi'_a(\xi)]^2} \quad \text{and} \quad \Psi''_a(\psi_a(\xi)) = -\frac{\psi''_a(\xi)}{[\psi'_a(\xi)]^3}. \quad (26)$$

Therefore, combining Eqs. (25) and (26), we get

$$\rho''_a(\psi_a(\xi)) = \frac{z''_a(\xi)}{[\psi'_a(\xi)]^2} - \frac{\psi''_a(\xi) \cdot z'_a(\xi)}{[\psi'_a(\xi)]^3} = \frac{z''_a(\xi) \cdot \psi'_a(\xi) - \psi''_a(\xi) \cdot z'_a(\xi)}{[\psi'_a(\xi)]^3},$$

as claimed. □

Lemma 5 For any integer $d > 0$ and any $\xi, a > 0$, we have

$$z'_a(\xi) = \frac{\tilde{g}'(\xi)\psi''_a(\xi) + \tilde{g}(\xi)\psi'_a(\xi)}{[\psi'_a(\xi)]^2}, \quad (27)$$

$$z''_a(\xi) = \frac{2[\psi''_a(\xi)]^2 \tilde{g}(\xi) - \psi'_a(\xi)\psi'''_a(\xi) \tilde{g}(\xi) - 2\tilde{g}'(\xi)\psi'_a(\xi)\psi''_a(\xi) + \tilde{g}''(\xi)[\psi'_a(\xi)]^2}{[\psi'_a(\xi)]^3}, \quad (28)$$

where

$$\tilde{g}(\xi) := \xi^{(d-2)/2} g(\xi) \quad \text{and} \quad \tilde{g}'(\xi) := \xi^{(d-2)/2} g'(\xi) + \frac{d-2}{2} \xi^{(d-4)/2} g(\xi).$$

Besides, we have that

$$\tilde{g}''(\xi) := (d-2)g'(\xi)\xi^{(d-4)/2} + g''(\xi)\xi^{(d-2)/2} + \frac{(d-2)(d-4)}{4} g(\xi)\xi^{(d-6)/2}.$$

Proof By Eq. (12), $z_a(\xi) := g(\xi)\xi^{(d-2)/2}/\psi'_a(\xi) = \tilde{g}(\xi)/\psi'_a(\xi)$ for $\xi > 0$. Differentiating this equality gives

$$z'_a(\xi) = \frac{\tilde{g}'(\xi)\psi'_a(\xi) - \tilde{g}(\xi)\psi''_a(\xi)}{[\psi'_a(\xi)]^2}.$$

We differentiate this expression a second time to obtain

$$z''_a(\xi) = \frac{(\tilde{g}'(\xi)\psi'_a(\xi) - \tilde{g}(\xi)\psi''_a(\xi))' [\psi'_a(\xi)]^2 - 2\psi'_a(\xi)\psi''_a(\xi)(\tilde{g}'(\xi)\psi'_a(\xi) - \tilde{g}(\xi)\psi''_a(\xi))}{[\psi'_a(\xi)]^4} - \frac{\tilde{g}(\xi)\psi''_a(\xi)}{[\psi'_a(\xi)]^4}$$

$$\begin{aligned}
&= \frac{(\tilde{g}'(\xi)\psi'_a(\xi) - \tilde{g}(\xi)\psi''_a(\xi))'\psi'_a(\xi) - 2\psi''_a(\xi)(\tilde{g}'(\xi)\psi'_a(\xi) - \tilde{g}(\xi)\psi''_a(\xi))}{[\psi'_a(\xi)]^3} \\
&= \frac{(\tilde{g}'(\xi)\psi''_a(\xi) + \tilde{g}''(\xi)\psi'_a(\xi) - \tilde{g}'(\xi)\psi''_a(\xi) - \tilde{g}(\xi)\psi'''_a(\xi))\psi'_a(\xi)}{[\psi'_a(\xi)]^3} \\
&\quad - \frac{2\psi''_a(\xi)(\tilde{g}'(\xi)\psi'_a(\xi) - \tilde{g}(\xi)\psi''_a(\xi))}{[\psi'_a(\xi)]^3} \\
&= \frac{(\tilde{g}''(\xi)\psi'_a(\xi) - \tilde{g}(\xi)\psi'''_a(\xi))\psi'_a(\xi) - 2\psi''_a(\xi)(\tilde{g}'(\xi)\psi'_a(\xi) - \tilde{g}(\xi)\psi''_a(\xi))}{[\psi'_a(\xi)]^3} \\
&= \frac{\tilde{g}''(\xi)[\psi'_a(\xi)]^2 - \tilde{g}(\xi)\psi'''_a(\xi)\psi'_a(\xi) - 2\tilde{g}'(\xi)\psi''_a(\xi)\psi'_a(\xi)}{[\psi'_a(\xi)]^3} \\
&\quad + \frac{2\tilde{g}(\xi)[\psi''_a(\xi)]^2}{[\psi'_a(\xi)]^3},
\end{aligned}$$

as claimed. □

Lemma 6 *We have*

$$\begin{aligned}
g''(\xi) &= \frac{\rho''_a(\psi_a(\xi))[\psi'_a(\xi)]^3}{\xi^{(d-2)/2}} \\
&\quad + \frac{\psi''_a(\xi) \times [\xi g(\xi) + 2\xi g'(\xi) + (d-2)g(\xi)] + g(\xi)\psi'''_a(\xi)}{\xi\psi'_a(\xi)} \\
&\quad + \left(\frac{\psi''_a(\xi)}{\psi'_a(\xi)}\right)^2 \left(g'(\xi) - 2g(\xi) + \frac{d-2}{2} \frac{g(\xi)}{\xi}\right) \\
&\quad - (d-2) \frac{g'(\xi)}{\xi} - \frac{(d-2)(d-4)}{4} \frac{g(\xi)}{\xi^2}.
\end{aligned}$$

Proof From Lemma 4 and Lemma 5, we have

$$\begin{aligned}
\rho''_a(\psi_a(\xi)) &= \frac{2\tilde{g}(\xi)[\psi''_a(\xi)]^2 - \tilde{g}(\xi)\psi'_a(\xi)\psi'''_a(\xi) - 2\tilde{g}'(\xi)\psi'_a(\xi)\psi''_a(\xi)}{[\psi'_a(\xi)]^5} \\
&\quad + \frac{\tilde{g}''(\xi)[\psi'_a(\xi)]^2 - \tilde{g}'(\xi)[\psi''_a(\xi)]^2 + \tilde{g}(\xi)\psi'_a(\xi)\psi''_a(\xi)}{[\psi'_a(\xi)]^5},
\end{aligned}$$

which yields

$$\begin{aligned}
\tilde{g}''(\xi) &= \frac{\rho_a''(\psi_a(\xi))[\psi_a'(\xi)]^5 + \tilde{g}'(\xi)[\psi_a''(\xi)]^2 + \tilde{g}(\xi)\psi_a'(\xi)\psi_a''(\xi)}{[\psi_a'(\xi)]^2} \\
&\quad + \frac{+\tilde{g}(\xi)\psi_a'(\xi)\psi_a'''(\xi) + 2\tilde{g}'(\xi)\psi_a'(\xi)\psi_a''(\xi) - 2\tilde{g}(\xi)[\psi_a''(\xi)]^2}{[\psi_a'(\xi)]^2} \\
&= \rho_a''(\psi_a(\xi))[\psi_a'(\xi)]^3 + \frac{\tilde{g}'(\xi)[\psi_a''(\xi)]^2}{[\psi_a'(\xi)]^2} + \frac{\tilde{g}(\xi)\psi_a''(\xi)}{\psi_a'(\xi)} \\
&\quad + \frac{\tilde{g}(\xi)\psi_a'''(\xi)}{\psi_a'(\xi)} + \frac{2\tilde{g}'(\xi)\psi_a''(\xi)}{\psi_a'(\xi)} - \frac{2\tilde{g}(\xi)[\psi_a''(\xi)]^2}{[\psi_a'(\xi)]^2} \\
&= \rho_a''(\psi_a(\xi))[\psi_a'(\xi)]^3 + \frac{\tilde{g}(\xi)\psi_a''(\xi) + \tilde{g}'(\xi)\psi_a''(\xi) + 2\tilde{g}'(\xi)\psi_a''(\xi)}{\psi_a'(\xi)} \\
&\quad + \frac{\tilde{g}'(\xi)[\psi_a''(\xi)]^2 - 2\tilde{g}(\xi)[\psi_a''(\xi)]^2}{[\psi_a'(\xi)]^2}.
\end{aligned}$$

Using the definition of $\tilde{g}''(\xi)$ given in Lemma 5, we express the above expression in terms of $g''(\xi)$. It yields that

$$\begin{aligned}
g''(\xi) &= \rho_a''(\psi_a(\xi))[\psi_a'(\xi)]^3 \\
&\quad + \frac{\xi^{(d-2)/2}g(\xi)\psi_a''(\xi) + \xi^{(d-2)/2}g(\xi)\psi_a'''(\xi) + 2\xi^{(d-2)/2}g'(\xi)\psi_a''(\xi)}{\psi_a'(\xi)} \\
&\quad + \frac{(d-2)\xi^{(d-4)/2}g(\xi)\psi_a''(\xi)}{\psi_a'(\xi)} \\
&\quad + \frac{[\xi^{(d-2)/2}g'(\xi) + \frac{d-2}{2}\xi^{(d-4)/2}g(\xi)][\psi_a''(\xi)]^2 - 2\xi^{(d-2)/2}g(\xi)[\psi_a''(\xi)]^2}{[\psi_a'(\xi)]^2} \\
&= \frac{1}{\xi^{(d-2)/2}} \times \left(\rho_a''(\psi_a(\xi))[\psi_a'(\xi)]^3 \right. \\
&\quad + \frac{\xi^{(d-4)/2}\psi_a''(\xi) \times [\xi g(\xi) + 2\xi g'(\xi) + (d-2)g(\xi)] + \xi^{(d-2)/2}g(\xi)\psi_a'''(\xi)}{\psi_a'(\xi)} \\
&\quad + \left(\frac{\psi_a''(\xi)}{\psi_a'(\xi)} \right)^2 \left(\xi^{(d-2)/2}[g'(\xi) - 2g(\xi)] + \frac{d-2}{2}\xi^{(d-4)/2}g(\xi) \right) \\
&\quad \left. - (d-2)g'(\xi)\xi^{(d-4)/2} - \frac{(d-2)(d-4)}{4}g(\xi)\xi^{(d-6)/2} \right) \\
&= \frac{\rho_a''(\psi_a(\xi))[\psi_a'(\xi)]^3}{\xi^{(d-2)/2}} + \frac{\psi_a''(\xi) \times [\xi g(\xi) + 2\xi g'(\xi) + (d-2)g(\xi)] + g(\xi)\psi_a'''(\xi)}{\xi\psi_a'(\xi)}
\end{aligned}$$

$$\begin{aligned}
& + \left(\frac{\psi_a''(\xi)}{\psi_a'(\xi)} \right)^2 \left(g'(\xi) - 2g(\xi) + \frac{d-2}{2} \frac{g(\xi)}{\xi} \right) \\
& - (d-2) \frac{g'(\xi)}{\xi} - \frac{(d-2)(d-4)}{4} \frac{g(\xi)}{\xi^2}.
\end{aligned}$$

□

Lemma 7 For any $\xi > 0$, we have

$$(z_a''\psi_a' - \psi_a''z_a')(\xi) = \tilde{g}''(\xi) + \frac{\tilde{C}_3 A^2 + \tilde{C}_4 A + \tilde{C}_5}{(\xi^{d/2} + A)^2},$$

where $A = A(a) := a^{d/2}$ and

$$\tilde{C}_3 := \frac{d-2}{4\xi^2} (3(d-2)\tilde{g}(\xi) - 6\tilde{g}'(\xi)\xi), \quad (29)$$

$$\tilde{C}_4 := \frac{d-2}{4\xi^2} (\tilde{g}(\xi)(d-4) - 6\tilde{g}'(\xi)\xi^{(d+2)/2}), \quad (30)$$

$$\tilde{C}_5 := \frac{d^2-4}{4\xi^2} \tilde{g}(\xi). \quad (31)$$

Proof We can express $z_a''\psi_a' - \psi_a''z_a'$ in terms of derivatives of ψ_a , by using (27) and (28). We have

$$\begin{aligned}
z_a''\psi_a' - \psi_a''z_a' &= \frac{2[\psi_a''(\xi)]^2 \tilde{g}(\xi) - \psi_a'(\xi)\psi_a'''(\xi)\tilde{g}(\xi) - 2\tilde{g}'(\xi)\psi_a'(\xi)\psi_a''(\xi)}{[\psi_a'(\xi)]^3} \psi_a' \\
&\quad + \frac{\tilde{g}''(\xi)[\psi_a'(\xi)]^2}{[\psi_a'(\xi)]^3} \psi_a' - \psi_a'' \frac{\tilde{g}'(\xi)\psi_a'(\xi) - \tilde{g}(\xi)\psi_a''(\xi)}{[\psi_a'(\xi)]^2} \\
&= \frac{3\tilde{g}(\xi)[\psi_a']^2 - \tilde{g}(\xi)\psi_a'\psi_a''' - 3\tilde{g}'(\xi)\psi_a'\psi_a'' + \tilde{g}''(\xi)[\psi_a']^2}{[\psi_a']^2}.
\end{aligned} \quad (32)$$

We now compute explicit expressions for the derivatives of ψ_a . Remember that $\psi_a(\xi) = -a + (A + \xi^{d/2})^{2/d}$ and $\psi_a'(\xi) = \xi^{\frac{d}{2}-1} (A + \xi^{d/2})^{\frac{2}{d}-1}$. Then, differentiating $\psi_a'(\xi)$ yields

$$\psi_a''(\xi) = \left(\frac{d}{2} - 1 \right) \xi^{\frac{d}{2}-2} \left(\xi^{\frac{d}{2}} + A \right)^{\frac{2}{d}-1} + \left(\frac{2}{d} - 1 \right) \xi^{\frac{d}{2}-1} \left(\xi^{\frac{d}{2}} + A \right)^{\frac{2}{d}-2} \frac{d}{2} \xi^{\frac{d}{2}-1}$$

$$\begin{aligned}
&= \left(\frac{d}{2} - 1\right) \xi^{\frac{d}{2}-2} \left(\xi^{\frac{d}{2}} + A\right)^{\frac{2}{d}-1} + \left(1 - \frac{d}{2}\right) \xi^{d-2} \left(\xi^{\frac{d}{2}} + A\right)^{\frac{2}{d}-2} \\
&= \frac{d-2}{2} \xi^{\frac{d}{2}-2} \left(\xi^{\frac{d}{2}} + A\right)^{\frac{2}{d}-2} \left(\left(\xi^{\frac{d}{2}} + A\right) - \xi^{\frac{d}{2}}\right) \\
&= \frac{d-2}{2} \xi^{\frac{d}{2}-2} A \left(\xi^{\frac{d}{2}} + A\right)^{\frac{2}{d}-2}.
\end{aligned}$$

This can be recognized as

$$\psi_a''(\xi) = \psi_a'(\xi) \times \frac{d-2}{2} \xi^{-1} A \left(\xi^{d/2} + A\right)^{-1}.$$

Differentiating again, we obtain

$$\begin{aligned}
\psi_a'''(\xi) &= \left(\frac{d-2}{2}\right) \left(\frac{d}{2} - 2\right) \xi^{\frac{d}{2}-3} A \left(\xi^{\frac{d}{2}} + A\right)^{\frac{2}{d}-2} \\
&\quad + \left(\frac{d-2}{2}\right) \left(\frac{2}{d} - 2\right) \xi^{\frac{d}{2}-2} A \left(\xi^{\frac{d}{2}} + A\right)^{\frac{2}{d}-3} \frac{d}{2} \xi^{\frac{d}{2}-1} \\
&= \frac{(d-2)(d-4)}{4} \xi^{\frac{d}{2}-3} A \left(\xi^{\frac{d}{2}} + A\right)^{\frac{2}{d}-2} \\
&\quad + \frac{(d-2)(1-d)}{2} \xi^{d-3} A \left(\xi^{\frac{d}{2}} + A\right)^{\frac{2}{d}-3} \\
&= \frac{d-2}{4} A \xi^{\frac{d}{2}-3} \left(\xi^{\frac{d}{2}} + A\right)^{\frac{2}{d}-3} \left((d-4) \left(\xi^{\frac{d}{2}} + A\right) + 2(1-d) \xi^{\frac{d}{2}}\right) \\
&= \frac{d-2}{4} A \xi^{\frac{d}{2}-3} \left(\xi^{\frac{d}{2}} + A\right)^{\frac{2}{d}-3} \left(d \xi^{\frac{d}{2}} + dA - 4 \xi^{\frac{d}{2}} - 4A + 2 \xi^{\frac{d}{2}} - 2d \xi^{\frac{d}{2}}\right) \\
&= \frac{d-2}{4} A \xi^{\frac{d}{2}-3} \left(\xi^{\frac{d}{2}} + A\right)^{\frac{2}{d}-3} \left(-d \xi^{\frac{d}{2}} - 2 \xi^{\frac{d}{2}} + dA - 4A\right) \\
&= \frac{d-2}{4} A \xi^{\frac{d}{2}-3} \left(\xi^{\frac{d}{2}} + A\right)^{\frac{2}{d}-3} \left((d-4)A - (d+2) \xi^{\frac{d}{2}}\right).
\end{aligned}$$

This can also be recognized as

$$\psi_a'''(\xi) = \psi_a'(\xi) \times \frac{d-2}{4} \xi^{-2} A \left(\xi^{d/2} + A\right)^{-2} \left((d-4)A - (d+2) \xi^{\frac{d}{2}}\right).$$

We have therefore shown that

$$\psi_a''(\xi) = \psi_a'(\xi) \times \frac{d-2}{2} \xi^{-1} A \left(\xi^{d/2} + A\right)^{-1},$$

$$\psi_a'''(\xi) = \psi_a'(\xi) \times \frac{d-2}{4} \xi^{-2} A \left(\xi^{d/2} + A \right)^{-2} \left((d-4)A - (d+2)\xi^{d/2} \right).$$

As a consequence, all terms in the numerator and the denominator of Eq. (32) are proportional to $[\psi_a']^2$, which can then be simplified, so that we obtain

$$\begin{aligned} z_a''\psi_a' - \psi_a''z_a' &= 3\tilde{g}(\xi) \frac{[\psi_a'']^2}{[\psi_a']^2} - \tilde{g}(\xi) \frac{\psi_a'\psi_a'''}{[\psi_a']^2} - 3\tilde{g}'(\xi) \frac{\psi_a'\psi_a''}{[\psi_a']^2} + \tilde{g}''(\xi) \\ &= \tilde{g}''(\xi) + \frac{d-2}{4\xi^2(\xi^{d/2} + A)^2} \times \\ &\quad \left(3(d-2)\tilde{g}(\xi)A^2 - \tilde{g}(\xi)((d-4)A - (d+2)\xi^{d/2}) \right. \\ &\quad \left. - 6\tilde{g}'(\xi)\xi A(\xi^{d/2} + A) \right) \\ &= \tilde{g}''(\xi) + \frac{\tilde{C}_3A^2 + \tilde{C}_4A + \tilde{C}_5}{(\xi^{d/2} + A)^2}, \end{aligned}$$

finishing the proof. \square

In the following lemma, we give an expression of $2\tilde{C}_3\xi^{d/2} - \tilde{C}_4$, $\tilde{C}_4\xi^{d/2} - 2\tilde{C}_5$, and $2\tilde{C}_5 + 2\tilde{C}_3\xi^d - 6\tilde{C}_4\xi^{d/2}$. These are conditions for the locations of the extrema of (24) in section “[Appendix 3: Proof of Proposition 4.2](#)”.

Lemma 8 *Let $\xi > 0$ and $d > 2$. Let \tilde{C}_3 , \tilde{C}_4 , and \tilde{C}_5 be defined as in Proposition 2.2. Then we have*

$$\begin{aligned} 2\tilde{C}_3\xi^{d/2} - \tilde{C}_4 &= \frac{d-2}{4\xi^2} \left(3\xi^{d-1}(d-2)g(\xi) - g(\xi)(d-4)\xi^{(d-2)/2} - 6g'(\xi)\xi^d \right) \\ 2\tilde{C}_5 - \tilde{C}_4\xi^{d/2} &= \frac{d-2}{4\xi^2} \left(2(d+2)\xi^{(d-2)/2}g(\xi) - \xi^{(d-2)/2}g(\xi)(d-4)\xi^{(d-2)/2} \right. \\ &\quad \left. + 3\xi^{(d-2)/2}(2g'(\xi) + (d-2)\xi^{-1}g(\xi))\xi^d \right). \end{aligned}$$

Lastly,

$$\begin{aligned} 2\tilde{C}_5 + 2\tilde{C}_3\xi^d - 6\tilde{C}_4\xi^{d/2} &= \frac{d-2}{4\xi^2} \left(2(d+2)\xi^{(d-2)/2}g(\xi) + 6(d-2)\xi^{(3d-2)/2}g(\xi) \right. \\ &\quad \left. - 3(2g'(\xi) + (d-2)\xi^{-1}g(\xi))(\xi^{(2d-2)/2}(d-4) - 4\xi^{3d/2}) \right). \end{aligned}$$

Proof We have

$$\begin{aligned}
2C_3C_6 - C_4 &= 2\xi^{d/2} \frac{d-2}{4\xi^2} \left(3(d-2)\tilde{g}(\xi) - 6\tilde{g}'(\xi)\xi \right) - \frac{d-2}{4\xi^2} \left(\tilde{g}(\xi)(d-4) \right. \\
&\quad \left. - 6\tilde{g}'(\xi)\xi^{(d+2)/2} \right) \\
&= \frac{d-2}{4\xi^2} \left(2\xi^{d/2} \left(3(d-2)\tilde{g}(\xi) - 6\tilde{g}'(\xi)\xi \right) - \left(\tilde{g}(\xi)(d-4) \right. \right. \\
&\quad \left. \left. - 6\tilde{g}'(\xi)\xi^{(d+2)/2} \right) \right) \\
&= \frac{d-2}{4\xi^2} \left(6\xi^{d/2}(d-2)\tilde{g}(\xi) - \tilde{g}(\xi)(d-4) - 6\tilde{g}'(\xi)\xi^{(d+2)/2} \right).
\end{aligned}$$

From Lemma 4, we get the expressions $\tilde{g}(\xi) := \xi^{(d-2)/2}g(\xi)$ and $2\tilde{g}'(\xi) = \xi^{(d-2)/2}(2g'(\xi) + (d-2)\xi^{-1}g(\xi))$.

$$\begin{aligned}
2C_3C_6 - C_4 &= \frac{d-2}{4\xi^2} \xi^{(d-2)/2} \left(6\xi^{d/2}(d-2)g(\xi) - g(\xi)(d-4) \right. \\
&\quad \left. - 3(2g'(\xi) + (d-2)\xi^{-1}g(\xi))\xi^{(d+2)/2} \right) \\
&= \frac{d-2}{4\xi^2} \xi^{(d-2)/2} \left(3\xi^{d/2}(d-2)g(\xi) - g(\xi)(d-4) - 6g'(\xi)\xi^{(d+2)/2} \right) \\
&= \frac{d-2}{4\xi^2} \left(3\xi^{d-1}(d-2)g(\xi) - g(\xi)(d-4)\xi^{(d-2)/2} - 6g'(\xi)\xi^d \right).
\end{aligned}$$

Furthermore,

$$\begin{aligned}
2C_5 - C_4C_6 &= 2\frac{d-2}{4\xi^2}(d+2)\tilde{g}(\xi) - \frac{d-2}{4\xi^2} \left(\tilde{g}(\xi)(d-4) \right. \\
&\quad \left. - 6\tilde{g}'(\xi)\xi^{(d+2)/2} \right) \xi^{(d-2)/2} \\
&= \frac{d-2}{4\xi^2} \left(2(d+2)\tilde{g}(\xi) - \left(\tilde{g}(\xi)(d-4) - 6\tilde{g}'(\xi)\xi^{(d+2)/2} \right) \xi^{(d-2)/2} \right) \\
&= \frac{d-2}{4\xi^2} \left(2(d+2)\tilde{g}(\xi) - \tilde{g}(\xi)(d-4)\xi^{(d-2)/2} + 6\tilde{g}'(\xi)\xi^d \right).
\end{aligned}$$

From Lemma 4, we get the expressions $\tilde{g}(\xi) := \xi^{(d-2)/2}g(\xi)$ and $2\tilde{g}'(\xi) = \xi^{(d-2)/2}(2g'(\xi) + (d-2)\xi^{-1}g(\xi))$.

$$\begin{aligned}
2C_5 - C_4C_6 &= \frac{d-2}{4\xi^2} \left(2(d+2)\xi^{(d-2)/2}g(\xi) - \xi^{(d-2)/2}g(\xi)(d-4)\xi^{(d-2)/2} \right. \\
&\quad \left. + 3\xi^{(d-2)/2}(2g'(\xi) + (d-2)\xi^{-1}g(\xi))\xi^d \right).
\end{aligned}$$

Using the fact that $C_6 = \xi^{d/2}$, we get

$$\begin{aligned}
& 2C_5 + 2C_3C_6^2 - 6C_4C_6 \\
&= 2C_5 + 2C_3\xi^d - 6C_4\xi^{d/2} \\
&= \frac{d-2}{4\xi^2} \left(2(d+2)\tilde{g}(\xi) + 2\xi^d(3(d-2)\tilde{g}(\xi) - 6\tilde{g}'(\xi)\xi) - 6\xi^{d/2}(\tilde{g}(\xi)(d-4) \right. \\
&\quad \left. - 6\tilde{g}'(\xi)\xi^{(d+2)/2}) \right) \\
&= \frac{d-2}{4\xi^2} \left(2(d+2)\tilde{g}(\xi) + 6\xi^d(d-2)\tilde{g}(\xi) - 12\tilde{g}'(\xi)\xi^{d+1} - 6\xi^{d/2}\tilde{g}(\xi)(d-4) \right. \\
&\quad \left. + 36\tilde{g}'(\xi)\xi^{(2d+2)/2} \right) \\
&= \frac{d-2}{4\xi^2} \left(2(d+2)\xi^{(d-2)/2}g(\xi) + 6(d-2)\xi^{(3d-2)/2}g(\xi) \right. \\
&\quad \left. - 6\xi^{(d-2)/2}(2g'(\xi) + (d-2)\xi^{-1}g(\xi))\xi^{d+1} \right. \\
&\quad \left. - 3\xi^{d/2}\xi^{(d-2)/2}(2g'(\xi) + (d-2)\xi^{-1}g(\xi))(d-4) \right. \\
&\quad \left. + 18\xi^{(d-2)/2}(2g'(\xi) + (d-2)\xi^{-1}g(\xi))\xi^{(2d+2)/2} \right) \\
&= \frac{d-2}{4\xi^2} \left(2(d+2)\xi^{(d-2)/2}g(\xi) + 6(d-2)\xi^{(3d-2)/2}g(\xi) \right. \\
&\quad \left. - 6\xi^{3d/2}(2g'(\xi) + (d-2)\xi^{-1}g(\xi)) \right. \\
&\quad \left. - 3\xi^{(2d-2)/2}(2g'(\xi) + (d-2)\xi^{-1}g(\xi))(d-4) \right. \\
&\quad \left. + 18\xi^{3d/2}(2g'(\xi) + (d-2)\xi^{-1}g(\xi)) \right).
\end{aligned}$$

So,

$$\begin{aligned}
& 2C_5 + 2C_3C_6^2 - 6C_4C_6 \\
&= \frac{d-2}{4\xi^2} \left(2(d+2)\xi^{(d-2)/2}g(\xi) + 6(d-2)\xi^{(3d-2)/2}g(\xi) \right. \\
&\quad \left. - 6\xi^{3d/2}(2g'(\xi) + (d-2)\xi^{-1}g(\xi)) \right. \\
&\quad \left. - 3\xi^{(2d-2)/2}(2g'(\xi) + (d-2)\xi^{-1}g(\xi))(d-4) \right)
\end{aligned}$$

$$\begin{aligned}
& + 18\xi^{3d/2}(2g'(\xi) + (d-2)\xi^{-1}g(\xi)) \Big) \\
& = \frac{d-2}{4\xi^2} \left(2(d+2)\xi^{(d-2)/2}g(\xi) + 6(d-2)\xi^{(3d-2)/2}g(\xi) \right. \\
& \quad \left. - 3(2g'(\xi) + (d-2)\xi^{-1}g(\xi))(2\xi^{3d/2} + \xi^{(2d-2)/2}(d-4) - 6\xi^{3d/2}) \right) \\
& = \frac{d-2}{4\xi^2} \left(2(d+2)\xi^{(d-2)/2}g(\xi) + 6(d-2)\xi^{(3d-2)/2}g(\xi) \right. \\
& \quad \left. - 3(2g'(\xi) + (d-2)\xi^{-1}g(\xi))(\xi^{(2d-2)/2}(d-4) - 4\xi^{3d/2}) \right).
\end{aligned}$$

□

Proof of Lemma 1

Proof Let $\Psi = \psi^{-1}$, i.e., $\Psi(x) = ((x+a)^{d/2} - a^{d/2})^{2/d}$. By Faà di Bruno's formula, we have

$$\rho^{(k)} = (z \circ \Psi)^{(k)} = \sum_{m=1}^k (z^{(m)} \circ \Psi) \times B_{k,m}(\Psi', \Psi'', \dots, \Psi^{(k-m+1)}),$$

where $B_{m,k}$ are the partial exponential Bell polynomials (see, e.g., [13] and the references therein). Therefore,

$$\rho^{(k)} \circ \psi = \sum_{m=1}^k z^{(m)} \times \tilde{B}_{k,m},$$

where

$$\tilde{B}_{k,m} := B_{k,m}(\Psi', \Psi'', \dots, \Psi^{(k-m+1)}) \circ \psi.$$

Remark that $z(\xi) = g(\xi) \times \tau(\xi)$, where $\tau(\xi) := \xi^{(d-2)/2}/\psi'(\xi)$. Therefore, applying the general Leibniz rule, we obtain

$$\rho^{(k)} \circ \psi = \sum_{m=1}^k (g \times \tau)^{(m)} \times \tilde{B}_{k,m} = \sum_{m=1}^k \sum_{i=0}^m \binom{m}{i} g^{(i)} \tau^{(m-i)} \tilde{B}_{k,m} = \sum_{i=0}^k \alpha_{i,k} g^{(i)},$$

where

$$\alpha_{0,k} := \sum_{m=1}^k \tau^{(m)} \widetilde{B}_{k,m}, \quad \text{and for } i \geq 1, \quad \alpha_{i,k} := \sum_{m=i}^k \binom{m}{i} \tau^{(m-i)} \widetilde{B}_{k,m}. \quad (33)$$

□

Applying again Faà di Bruno's formula, it is possible to compute higher order derivatives of Ψ and τ . Indeed, $\Psi(x) = ((x+a)^{d/2} - a^{d/2})^{2/d} = f_1 \circ f_2$, where $f_1(x) = x^{2/d}$ and $f_2(x) = (x+a)^{d/2} - a^{d/2}$. Therefore,

$$\Psi^{(k)} = \sum_{m=1}^k (f_1^{(m)} \circ f_2) \times B_{k,m}(f_2', f_2'', \dots, f_2^{(k-m+1)}).$$

Note that for $m \geq 1$,

$$f_1^{(m)}(x) = \frac{\prod_{i=0}^{m-1} (2-i \times d)}{d^m} x^{(2/d)-m}, \quad \text{and}$$

$$f_2^{(m)}(x) = \frac{\prod_{i=0}^{m-1} (d-2i)}{2^m} (x+a)^{(d-2m)/2}.$$

We have $\tau(\xi) := \xi^{(d-2)/2} / \psi'(\xi)$. Remember that $\psi'_a(\xi) = \xi^{\frac{d}{2}-1} (A + \xi^{d/2})^{\frac{2}{d}-1}$. Therefore, $\tau(\xi) = (A + \xi^{d/2})^{(d-2)/d} = f_3 \circ f_4$, where $f_3(x) = A + x^{d/2}$ and $f_4(x) = x^{(d-2)/d}$. The computation of the derivatives of τ can be done in the same way as for Ψ .

Appendix 5: Proof of Proposition 3.1: MSE for Estimating Derivatives

Computation of the Bias

Proof To simplify the notations, we write only $\widehat{\eta}$ instead of $\widehat{\eta}_{k,n,h,a}(\xi)$. By linearity, we have

$$\mathbb{E}[\widehat{\eta}] = \frac{1}{s_d h^{k+1}} \left(\mathbb{E} \left[K_k \left(\frac{\psi_a(\xi) - \psi_a(\xi_1)}{h} \right) \right] + \mathbb{E} \left[K_k \left(\frac{\psi_a(\xi) + \psi_a(\xi_1)}{h} \right) \right] \right).$$

By assumption, the kernel K_k has a compact support; we can then apply Lemma 2 and obtain

$$K_k \left(\frac{\psi_a(\xi) + \psi_a(\xi_1)}{h} \right) = 0,$$

for n large enough. Therefore, by applying Stute and Werner's [15] change of variable

$$\begin{aligned}\mathbb{E}[\widehat{\eta}] &= \frac{1}{s_d h^{k+1}} \int_{\mathbb{R}^d} K_k \left(\frac{\psi_a(\xi) - \psi_a(\xi_1)}{h} \right) g(\xi_1) |\Sigma|^{-1/2} d\xi_1 \\ &= \frac{s_d}{s_d h^{k+1}} \int_0^\infty K_k \left(\frac{\psi_a(\xi) - \psi_a(v_1)}{h} \right) g(v_1) v_1^{(d-2)/2} dv_1.\end{aligned}$$

We now do a second change of variables

$$w := w(v_1) = \frac{-\psi_a(\xi) + \psi_a(v_1)}{h} \iff v_1(w) = \psi_a^{-1}(\psi_a(\xi) + wh),$$

which gives

$$\mathbb{E}[\widehat{\eta}] = \frac{1}{h^{k+1}} h \int_{-\psi_a(\xi)/h}^\infty K_k(-w) \frac{g(\psi_a^{-1}(\psi_a(\xi) + wh)) \cdot v_1(w)^{(d-2)/2}}{\psi_a'(\psi_a^{-1}(\psi_a(\xi) + wh))} dw.$$

We get that

$$\begin{aligned}\mathbb{E}[\widehat{\eta}] &= \frac{1}{h^k} \int_{-\psi_a(\xi)/h}^\infty K_k(-w) z(\psi_a^{-1}(\psi_a(\xi) + wh)) dw \\ &= \frac{1}{h^k} \int_{-\infty}^\infty K_k(-w) \cdot z(\psi_a^{-1}(\psi_a(\xi) + wh)) dw \\ &\quad - \frac{1}{h^k} \int_{-\infty}^{-\psi_a(\xi)/h} K_k(-w) \cdot z(\psi_a^{-1}(\psi_a(\xi) + wh)) dw.\end{aligned}$$

We have $(-\psi_a(\xi))/h \rightarrow -\infty$ as $n \rightarrow +\infty$. Since K_k has compact support, we deduce that the second term in the above equation is zero for n large enough by Lemma 2. We now then have

$$\begin{aligned}\mathbb{E}[\widehat{\eta}] &= \frac{1}{h^k} \int_{-\infty}^\infty K_k(-w) \cdot z(\psi_a^{-1}(\psi_a(\xi) + wh)) dw \\ &= \frac{1}{h^k} \int_{-\infty}^\infty K_k(-w) \cdot \rho_a(\psi_a(\xi) + wh) dw.\end{aligned}$$

We now apply the Taylor-Lagrange expansion at the order $m+1$ to the function ρ_a at the point $\psi_a(\xi)$.

$$\mathbb{E}[\widehat{\eta}] = \int_{-\infty}^\infty K_k(-w) \cdot z(\psi_a^{-1}(\psi_a(\xi) + wh)) dw$$

$$\begin{aligned}
&= \sum_{i=0}^m \int_{-\infty}^{\infty} K_k(-w) \cdot w^i \frac{h^i}{i!} \rho_a^{(i)}(\psi_a(\xi)) dw \\
&\quad + \int_{-\infty}^{\infty} K_k(-w) \cdot w^{m+1} \frac{h^{m+1}}{(m+1)!} \rho_a^{(m+1)}(c_T) dw,
\end{aligned}$$

where c_T lies between $\psi_a(\xi) + wh$ and $\psi_a(\xi)$. All the terms in the sum are zero except for $i = k$ and $i = m$. Therefore,

$$\begin{aligned}
\mathbb{E}[\widehat{\eta}] &= \int_{-\infty}^{\infty} K_k(-w) \cdot w^k \frac{h^k}{k!} \rho_a^{(k)}(\psi_a(\xi)) dw \\
&\quad + \int_{-\infty}^{\infty} K_k(-w) \cdot w^m \frac{h^m}{m!} \rho_a^{(m)}(\psi_a(\xi)) dw \\
&\quad + \int_{-\infty}^{\infty} K_k(-w) \cdot w^{m+1} \frac{h^{m+1}}{(m+1)!} \rho_a^{(m+1)}(c_T) dw.
\end{aligned}$$

By the assumption on the kernel, we get

$$\begin{aligned}
\mathbb{E}[\widehat{\eta}] &= \rho_a^{(k)}(\psi_a(\xi)) + \mu_m(K_k) \frac{h^{m-k}}{m!} \rho_a^{(m)}(\psi_a(\xi)) \\
&\quad + \frac{h^{m-k+1}}{(m+1)!} \int_{-\infty}^{\infty} K_k(-w) \cdot w^{m+1} \rho_a^{(m+1)}(c_T) dw.
\end{aligned}$$

If $\rho_a^{(m+1)}$ is continuous, then $\rho_a^{(m+1)}$ is bounded around $\psi_a(\xi)$. Therefore, the last integral will converge to a finite constant by Bochner's lemma. So,

$$\mathbb{E}[\widehat{\eta}] = \rho_a^{(k)}(\psi_a(\xi)) + \mu_m(K_k) \frac{h^{m-k}}{m!} \rho_a^{(m)}(\psi_a(\xi)) + O(h^{m-k+1}).$$

The bias of $\widehat{\eta}$ is

$$\text{Bias}_{\widehat{\eta}}(\xi) = \mu_m(K_k) \frac{h^{m-k}}{m!} \rho_a^{(m)}(\psi_a(\xi)) + O(h^{m-k+1}).$$

□

Variance for Estimating Derivatives and the Corresponding MSE

We now study the variance term. From the definition of the variance, we have

$$\text{Var}(\widehat{\eta}) = \mathbb{E}[\widehat{\eta}^2] - \mathbb{E}[\widehat{\eta}]^2.$$

Since that the term $\mathbb{E}[\widehat{\eta}]^2$ is already known, we only need to study the term $\mathbb{E}[\widehat{\eta}^2]$. We have

$$\begin{aligned} \mathbb{E}[\widehat{\eta}^2] &= \frac{\Gamma(d/2)^2}{\pi^d n^2 h^{2(k+1)}} \mathbb{E} \left[\left(\sum_{i=1}^n \left[K_k \left(\frac{\psi_a(\xi) - \psi_a(\xi_i)}{h} \right) \right. \right. \right. \\ &\quad \left. \left. \left. + K_k \left(\frac{\psi_a(\xi) + \psi_a(\xi_i)}{h} \right) \right] \right)^2 \right]. \end{aligned}$$

Expanding the sum, we find that

$$\begin{aligned} \mathbb{E}[\widehat{\eta}^2] &= \frac{\Gamma(d/2)^2}{\pi^d n^2 h^{2(k+1)}} \left(\sum_{i=1}^n (T_{1,i} + T_{2,i} + 2T_{3,i}) \right. \\ &\quad \left. + \sum_{1 \leq i \neq j \leq n} (T_{4,i,j} + T_{5,i,j} + T_{6,i,j} + T_{7,i,j}) \right), \end{aligned}$$

where

$$\begin{aligned} T_{1,i} &:= \mathbb{E} \left[K_k^2 \left(\frac{\psi_a(\xi) - \psi_a(\xi_i)}{h} \right) \right] \\ T_{2,i} &:= \mathbb{E} \left[K_k^2 \left(\frac{\psi_a(\xi) + \psi_a(\xi_i)}{h} \right) \right] \\ T_{3,i} &:= \mathbb{E} \left[K_k \left(\frac{\psi_a(\xi) - \psi_a(\xi_i)}{h} \right) K_k \left(\frac{\psi_a(\xi) + \psi_a(\xi_i)}{h} \right) \right] \\ T_{4,i,j} &:= \mathbb{E} \left[K_k \left(\frac{\psi_a(\xi) - \psi_a(\xi_i)}{h} \right) K_k \left(\frac{\psi_a(\xi) - \psi_a(\xi_j)}{h} \right) \right] \\ T_{5,i,j} &:= \mathbb{E} \left[K_k \left(\frac{\psi_a(\xi) - \psi_a(\xi_i)}{h} \right) K_k \left(\frac{\psi_a(\xi) + \psi_a(\xi_j)}{h} \right) \right] \\ T_{6,i,j} &:= \mathbb{E} \left[K_k \left(\frac{\psi_a(\xi) + \psi_a(\xi_i)}{h} \right) K_k \left(\frac{\psi_a(\xi) - \psi_a(\xi_j)}{h} \right) \right] \\ T_{7,i,j} &:= \mathbb{E} \left[K_k \left(\frac{\psi_a(\xi) + \psi_a(\xi_i)}{h} \right) K_k \left(\frac{\psi_a(\xi) + \psi_a(\xi_j)}{h} \right) \right]. \end{aligned}$$

Therefore, because the random variables $\mathbf{X}_1, \dots, \mathbf{X}_n$ are identically distributed, we get

$$\mathbb{E} \left[\widehat{\eta}^2 \right] = \frac{\Gamma(d/2)^2}{\pi^d n h^{2(k+1)}} (T_{1,1} + T_{2,1} + 2T_{3,1} + (n-1)T_{4,1,2} + (n-1)T_{5,1,2} \\ + (n-1)T_{6,1,2} + (n-1)T_{7,1,2}).$$

Step 1. Removing Terms That Disappear We start by applying Lemma 2. This means that for all n large enough,

$$\mathbb{E} \left[\widehat{\eta}^2 \right] = \frac{\Gamma(d/2)^2}{\pi^d n h^{2(k+1)}} (T_{1,1} + (n-1)T_{4,1,2}).$$

Step 2. Computation of an Equivalent of $T_{1,1}$ We apply Lemma 3, with K^2 instead of K , and obtain

$$T_{1,1} = \frac{\pi^{d/2} h}{\Gamma(d/2)} \int_{-\psi_a(\xi)/h}^{\infty} K_k^2(-w) \frac{g(\psi_a^{-1}(\psi_a(\xi) + wh)) \cdot v_1(w)^{(d-2)/2}}{\psi_a'(\psi_a^{-1}(\psi_a(\xi) + wh))} dw$$

and so

$$\frac{\Gamma(d/2)^2}{\pi^d n h^{2(k+1)}} T_{1,1} = \frac{\Gamma(d/2)}{\pi^{d/2}} \frac{1}{n h^{2k+1}} \\ \times \int_{-\psi_a(\xi)/h}^{\infty} K_k^2(-w) \frac{g(\psi_a^{-1}(\psi_a(\xi) + wh)) \cdot v_1(w)^{(d-2)/2}}{\psi_a'(\psi_a^{-1}(\psi_a(\xi) + wh))} dw \\ \sim \frac{\Gamma(d/2)}{\pi^{d/2}} \frac{1}{n h^{2k+1}} g(\xi) \int_{-\infty}^{\infty} K_k^2(w) dw,$$

as $n \rightarrow \infty$, by Bochner's lemma and remarking that $\int_{-\infty}^{\infty} K_k^2(w) dw = \int_{-\infty}^{\infty} K_k^2(-w) dw$.

Step 3. Computation of $T_{4,1,2}$ Since $\mathbf{X}_1, \dots, \mathbf{X}_n$ are independent and identically distributed, we have

$$T_{4,1,2} = \mathbb{E} \left[K_k \left(\frac{\psi_a(\xi) - \psi_a(\xi_1)}{h} \right) \right] \cdot \mathbb{E} \left[K_k \left(\frac{\psi_a(\xi) - \psi_a(\xi_2)}{h} \right) \right] \\ = \mathbb{E} \left[K_k \left(\frac{\psi_a(\xi) - \psi_a(\xi_1)}{h} \right) \right]^2 \\ = \frac{\pi^d h^{2(k+1)}}{\Gamma(d/2)^2} \mathbb{E} [\widehat{\eta}]^2,$$

for n large enough by Lemma 2.

Step 4. Combining All Previous Results

$$\begin{aligned} \text{Var}(\widehat{\eta}) &= \mathbb{E}[\widehat{\eta}^2] - \mathbb{E}[\widehat{\eta}]^2 \\ &= \frac{\Gamma(d/2)^2}{\pi^d n h^{2(k+1)}} (T_{1,1} + (n-1)T_{4,1,2}) - \mathbb{E}[\widehat{g}_{n,h,a}(\mathbf{x})]^2 \\ &= \frac{\Gamma(d/2)^2}{\pi^d n h^{2(k+1)}} (T_{1,1} + (n-1)T_{4,1,2}) - \left(\rho_a^{(k)}(\psi_a(\xi)) - \text{Bias}_{\widehat{\eta}}(\mathbf{x})\right)^2 \\ &= \frac{n-1}{n} \left(\rho_a^{(k)}(\psi_a(\xi)) - \text{Bias}_{\widehat{\eta}}(\mathbf{x})\right)^2 + \frac{\Gamma(d/2)^2}{\pi^d n h^{2(k+1)}} T_{1,1} \\ &\quad - \left(\rho_a^{(k)}(\psi_a(\xi)) - \text{Bias}_{\widehat{\eta}}(\mathbf{x})\right)^2 \\ &= -\frac{1}{n} \left(\rho_a^{(k)}(\psi_a(\xi)) - \text{Bias}_{\widehat{\eta}}(\mathbf{x})\right)^2 + \frac{\Gamma(d/2)^2}{\pi^d n h^{2(k+1)}} T_{1,1}. \end{aligned}$$

Since $\text{Bias}_{\widehat{\eta}}(\mathbf{x}) = o(1)$, we get

$$\begin{aligned} \text{Var}(\widehat{\eta}) &= \frac{\Gamma(d/2)^2}{\pi^d n h^{2(k+1)}} T_{1,1} + O(1/n) \\ &= \frac{\Gamma(d/2)}{\pi^{d/2}} \frac{1}{n h^{2k+1}} g(\xi) \|K_k\|_2^2 + o(n^{-1} h^{-(2k+1)}) + O(n^{-1}) \\ &\sim \frac{\Gamma(d/2)}{\pi^{d/2}} \frac{1}{n h^{2k+1}} g(\xi) \|K_k\|_2^2, \end{aligned}$$

since $h(n) \rightarrow 0$ as $n \rightarrow \infty$. So,

$$\begin{aligned} \text{MSE}[\widehat{\eta}] &= \left(\mu_m(K_k) \frac{h^{m-k}}{m!} \rho_a^{(m)}(\psi_a(\xi))\right)^2 + \frac{\Gamma(d/2)}{\pi^{d/2}} \frac{1}{n h^{2k+1}} g(\xi) \|K_k\|_2^2 \\ &\quad + O(h^{2(m-k+1)}) + o(n^{-1} h^{-(2k+1)}) + O(n^{-1}). \end{aligned}$$

References

1. Battey, H., Linton, O.: Nonparametric estimation of multivariate elliptic densities via finite mixture sieves. *J. Multivariate Anal.* **123**, 43–67 (2014)
2. Bhattacharyya, S., Bickel, P.J.: Adaptive estimation in elliptical distributions with extensions to high dimensions (2012)
3. Delft High Performance Computing Centre (DHPC): DelftBlue Supercomputer (Phase 2) (2024). <https://www.tudelft.nl/dhpc/ark:/44463/DelftBluePhase2>
4. Derumigny, A., Fermanian, J.-D.: Identifiability and estimation of meta-elliptical copula generators. *J. Multivariate Anal.* **190**, 104962 (2022)

5. Derumigny, A., Fermanian, J.-D., Min, A.: Testing for equality between conditional copulas given discretized conditioning events. *Can. J. Stat.* **51**(4), 1084–1110 (2023)
6. Derumigny, A., Fermanian, J.-D., Ryan, V.: *ElliptCopulas: Inference of Elliptical Distributions and Copulas* (2024). R package version 0.1.4. <https://cran.r-project.org/package=ElliptCopulas>, <https://github.com/AlexisDerumigny/ElliptCopulas>
7. Derumigny, A., Schmidt-Hieber, J.: On lower bounds for the bias-variance trade-off. *Ann. Stat.* **51**(4), 1510–1533 (2023)
8. Engel, J., Herrmann, E., Gasser, T.: An iterative bandwidth selector for kernel estimation of densities and their derivatives. *J. Nonparam. Stat.* **4**(1), 21–34 (1994)
9. Kai-Tai, F., Yao-Ting, Z.: *Generalized Multivariate Analysis*. Springer, Berlin (1990)
10. Liang, P.: *Nonparametric Estimation of Elliptical Copulas*. Ph.D. Thesis, University of Texas at El Paso, 2023
11. Liebscher, E.: A semiparametric density estimator based on elliptical distributions. *J. Multivariate Anal.* **92**(1), 205–225 (2005)
12. Maronna, R.A.: Robust M-estimators of multivariate location and scatter. *Ann. Stat.* 51–67 (1976)
13. Nardo, E.D., Guarino, G.: *kStatistics: Unbiased Estimators for Cumulant Products and Faa Di Bruno's Formula* (2022). R package version 2.1.1
14. Sadefo Kamdem, J.: Value-at-risk and expected shortfall for linear portfolios with elliptically distributed risk factors. *Int. J. Theor. Appl. Financ.* **8**, 537–551 (2005)
15. Stute, W., Werner, U.: Nonparametric estimation of elliptically contoured densities. In: *Nonparametric Functional Estimation and Related Topics*, pp. 173–190 (1991)
16. Tsybakov, A.B.: Nonparametric estimators. In: *Introduction to Nonparametric Estimation* (2009)