Confidence Bands in Functional Data Analysis based on the Kac-Rice Formula

Master Thesis

in the study program of Business Mathematics
at the Department of Mathematics and Computer Science,

Philipps-Universität Marburg

submitted by
Lisa Drescher
on May 10, 2024

supervised by
Prof. Dr. Hajo Holzmann

Contents

1. Introduction

Having confidence regions for parameters of observed data opens up a variety of possibilities compared to parameter estimation. A confidence region provides a measure for the uncertainty of the estimation as well as a range in which the true parameter is found with high probability, especially allowing to conduct hypothesis tests. This improves interpretability and thus also enhances utility of the data.
While the construction of confidence intervals for real-valued parameters is a well-studied field, we consider in this work confidence bands for functional parameters of random processes. This problem can be assigned to the field of functional data analysis (FDA).

The field of FDA gained growing interest over the past couple of decades, driven by increasing technical possibilities to collect and store large amounts of data, or even new ways to measure data. We may think about smart watches tracking health data of individuals, nearly continuously in time or data arising over time from climate or meteorology observations. These are just two examples of the types of data that can be collected and analyzed using FDA techniques. Ramsay and Silverman (2005) initiated the development of this field with the first version of their textbook on FDA, which appeared in 1997.
FDA involves the statistical methods related to data that are functional. This means that our observations are functions over a continuous domain. The domain must not necessarily be restricted to the univariate case, as considered in this work. Also, general random fields over subsets of $\mathbb{R}^n$ may be considered for which theory can be found in Adler and Taylor (2007).
Typically, however, observations are still made at a discrete grid of points. If the grid is sufficiently fine, i.e., contains a large number of observation points in relation to the sample size, it is referred to as dense. The design of the observation grid must be carefully considered, as its sparsity or density will influence not only the asymptotic properties observed but may also affect the choice of estimation procedures, c.f. Zhang and Wang (2016), Berger et al. (2023), Berger and Holzmann (2024).
Within the broad landscape of FDA, this work only revolves around univariate processes with continuously differentiable paths, observed on a fixed equidistant grid. Following Liebl and Reimherr (2023) as the primary source of this work, simultaneous confidence bands are constructed for univariate parameter functions of such processes.

There exist a number of methodologies for the construction of simultaneous confidence bands. For instance, Choi and Reimherr (2018) utilize the Karhunen–Loève expansion of an asymptotically normal estimator of the functional parameter in order to construct simultaneous confidence regions, thereby obtaining hyperellipsoids and hyper-rectangles as confidence regions for the functional parameter. They also propose an approach to transform confidence hyperellipsoids into valid confidence regions since procedures based on dimension reduction methods, as principal component analysis, often lead to confidence regions with zero-coverage.
Furthermore, numerous methods exist for constructing confidence bands based on resampling techniques, for example Degras (2011) propose bands based parametric bootstrapping, or refer to Neumann and Polzehl (1998).

Our approach to construct confidence bands for some functional parameter is based significantly on a suitable Kac-Rice formula which delivers the expected number of up-crossings of some random process above some deterministic functional bound.
Therefore, we consider the standardized version $\{X(t), t \in [0,1]\}$ of the functional parameter $\{\theta(t), t \in [0,1]\}$ which we assume to be elliptically distributed. For such processes $X(t)$, we present a respective Kac-Rice formula that calculates the expected number of up-crossings of $X(t)$ above a non-constant, adaptive critical value function $u(t)$, denoted by $\mathbb{E}[\varphi_{u,X}]$.

It turns out that $\mathbb{E}[\varphi_{u,X}]$ is an upper bound of the excursion probability $\mathbb{P}(\exists\, t \in [0,1] : X(t) \geq u(t))$ which is essentially considered for the construction of the simultaneous confidence bands. For obtaining a $(1-\alpha)$ band, we equate $\mathbb{E}[\varphi_{u,X}]$ with $\alpha/2$ and solve for the functional upper bound $u(t)$.

Furthermore, this Kac-Rice formula itself depends on some roughness parameter function $\tau(t)$ of the process $X(t)$. In contrast to the approach from Liebl and Reimherr (2023), we will propose a way to estimate the roughness function $\tau(t)$ that also works in the case of discrete observations with pointwise additive noise. In order to achieve this, we employ the bivariate local polynomial estimator of the covariance kernel, as in Berger and Holzmann (2024). This estimator effectively removes the noise and, in particular, delivers estimates of the derivatives of the covariance kernel such that the roughness parameter function can be derived.

Afterwards, we present a methodology for the construction of fair confidence bands in terms of balanced false positive rates across a partition of the domain. This includes a concrete algorithm which we newly introduce in order to simplify the one proposed by Liebl and Reimherr (2023) while still obtaining the same results.
Unlike Liebl and Reimherr (2023), who alternate correction locations between the left and right boundaries of partition intervals, our method consistently uses the left boundary. This adjustment allows us to employ a simpler version of the Kac-Rice formula, also making the computational process more straightforward.
We also quantify the conservatism, denoted by price of fairness, added to the confidence bands by incorporating the fairness constraint with the different algorithms. Our findings, including a correction of the price of fairness stated by Liebl and Reimherr (2023) for their algorithm, indicate that the complexity of the Liebl and Reimherr (2023) algorithm may not be necessary.

This work is structured as follows: The rather technical Section 2 commences with an introduction to the setting of this work, after which it presents the Kac-Rice formula in a simplified as well as more generalized form. We also provide some relevant special cases and then, conclude with a detailed and stepwise proof.
In Section 3, we explain how to construct confidence bands for any parameter function $\theta(t)$ using the Kac-Rice formula. Even though, generally, one needs to know the distribution of the standardized version $X(t)$, we also justify our construction with the Kac-Rice formula in cases where only the asymptotic distribution of $X(t)$ is known. Afterwards, we discuss the example of confidence bands for the mean function which will be the leading example in this work.
Further, for the Kac-Rice formula, the knowledge or consistent estimation of the roughness parameter function $\tau(t)$ of the underlying process $X(t)$ is essential. Thus, in Section 4, we describe two approaches for estimating $\tau(t)$: The first approach is that used by Liebl and Reimherr (2023), and the second is a new approach that employs the bivariate local polynomial estimation, see Berger and Holzmann (2024). For the latter approach, a brief bandwidth analysis is conducted and finally, both approaches are compared.
The concept of fairness is motivated and elaborated in Section 5. We specify the definition and our understanding of fairness, i.e. balanced false positive rates, and introduce a new, simpler algorithm for computing fair confidence bands. Moreover, we assess the price of fairness which is a quantity that measures the conservatism we obtain by constructing fair confidence bands.
Finally, Section 6 analyzes the empirical coverage probability of the confidence bands for the mean function.

2. Kac-Rice Formula

Formulas that calculate the expected number of up-crossings of a stochastic process above a deterministic critical value are grouped together as "Kac-Rice formulas". Originally introduced by Stephen Oswald Rice and Mark Kac, the Kac-Rice formula treated the case of a Gaussian process $X(t)$ and a constant upper bound $u(t) \equiv u$. This classic formulation can be found, e.g., in the book of Adler and Taylor (2007), Chapter 11.1.
In this section, we present a generalized version of the Kac-Rice formula, first introduced by Liebl and Reimherr (2023). The generalization compared to the classic Kac-Rice formula consists of allowing not only Gaussian processes and constant bounds but also elliptical random processes $X(t)$ and adaptive, non-constant critical value functions $u(t)$.

This formula plays a key role in the method we are going to use in order to construct confidence bands by providing an explicit expression for the expected number of times a random process $X(t)$ exceeds the deterministic critical value function $u(t)$. Due to the importance of the Kac-Rice formula for our construction of confidence bands, we present it at the beginning, even though its one of the more technical and theoretical parts of this work.
In the following, we will first clarify the framework and relevant terms as well as introduce necessary definitions. Particularly, we formulate assumptions on the stochastic process $X(t)$. Following that, a detailed proof of our Kac-Rice formula will be presented. Moreover, we will address the scenario where only the asymptotic distribution of the process X is known. Lastly, we will present some special cases of the formula that will be relevant for our practical purposes later in this work.

2.1. Setting and Kac-Rice Formula

We restrict the domain to the unit interval $[0, 1]$ which does not compromise generality and which is a common simplification of notation in the field of functional data analysis. Similarly to Liebl and Reimherr (2023), Assumption 3.1, or Adler and Taylor (2007), Chapter 11.2, we assume that the sample paths are almost surely continuously differentiable.

Assumption 2.1. We assume that $X = \{X(t), t \in [0, 1]\}$ is a centered elliptical process with $X \in C^1[0, 1]$ almost surely.

Definition 2.2 (**Covariance Function**)**.** The covariance function of a stochastic process $X(t)$ is defined as the bivariate function

$$C(t, s) := \mathrm{Cov}\big(X(t), X(s)\big) = \mathbb{E}\big[\big(X(t) - \mathbb{E}[X(t)]\big)\big(X(s) - \mathbb{E}[X(s)]\big)\big], \quad t, s \in [0, 1].$$

In our case the covariance operator does not have finite rank so that we can utilize the following characterization of elliptical processes proposed by Boente et al. (2014), saying that elliptical processes are stochastic processes that can be expressed as scalar mixtures of Gaussians.

Lemma 2.3 (**Characterization of Elliptical Processes**)**.** If $X = \{X(t), t \in [0, 1]\}$ is a centered elliptical process with a covariance operator that does not have finite rank, then there exists a strictly positive random variable $V > 0$, denoted by mixing coefficient, and a mean-zero Gaussian process $Z = \{Z(t), t \in [0, 1]\}$ such that V and Z are independent and satisfy $\{X(t), t \in [0, 1]\} \overset{d}{=} \{VZ(t), t \in [0, 1]\}$.

Moreover, when $\mathbb{E}[V^2] < \infty$, the covariance operator of X is[1]

$$\begin{aligned} \mathrm{Cov}\big(X(t), X(s)\big) &= \mathbb{E}[X(t)X(s)] \\ &= \mathbb{E}[VZ(t)VZ(s)] = \mathbb{E}[V^2]\mathrm{Cov}\big(Z(t), Z(s)\big), \quad t, s \in [0, 1]. \end{aligned} \tag{2.1}$$

[1] see Boente et al. (2014)

Note that $\mathrm{Cov}(Z(t), Z(s))$, $t, s \in [0,1]$, exists even when the second moment of X may not.

Further, like Liebl and Reimherr (2023), we consider critical value functions that fulfill the following.

Assumption 2.4. The critical value functions $u : [0,1] \to \mathbb{R}$ belong to $C^1_{\text{a.e.}}[0,1]$, meaning they are continuous across the entire domain and continuously differentiable almost everywhere.

Working towards understanding the Euler characteristic, we define the excursion set. In contrast to Adler and Taylor (2007), Definition 6.0.1, we allow a functional bound u.

Definition 2.5 (**Excursion Set**)**.** For two measurable functions $X, u : [0,1] \to \mathbb{R}$, we call

$$\{t \in [0,1] : X(t) \geq u(t)\} \tag{2.2}$$

the excursion set of $X(t)$ above the critical value function $u(t)$ here.

Note that the above definition works the same for a stochastic process $X(t)$ whose paths are in $C^1[0,1]$ as in Assumption 2.1 and thus measurable.

For the sake of well-definedness in the next Definition 2.7, we must first make one more assumption on the stochastic process X to not have any touch points with u or up-crossings with the same gradient as u. This is equivalent to the assumption on the process $X - u$ to not have any zeros of higher order.

Assumption 2.6. We assume for the stochastic process $X = \{X(t), t \in [0,1]\}$ with $X \in C^1[0,1]$ and the function $u = \{u(t), t \in [0,1]\}$ with $u \in C^1_{a.e.}[0,1]$ that there exists no point $t \in [0,1]$ such that

$$X(t) - u(t) = X'(t) - u'(t) = 0.$$

In the book of Adler and Taylor (2007), an equivalent requirement was made to classify a function as "suitably regular", see their Definition 6.2.1, for a similar purpose. Also Azaïs and Wschebor (2009) formulate this Assumption, see their Lemma 3.1, in the context of the Kac-Rice formula.

Definition 2.7 (**Up-Crossings**)**.** The number of up-crossings of $X(t)$ about $u(t)$ on the interval $[a,b]$, $0 \leq a < b \leq 1$ is defined by

$$N_{u,X}[a,b] = \#\{t \in [a,b] : X(t) = u(t),\, X'(t) > u'(t)\}.$$

Now, we consider the Euler characteristic of the excursion set (2.2) which is the number of disjoint intervals and isolated points of the set. For more details on Euler characteristics, see Adler and Taylor (2007), Theorem 6.1.1. Together with Assumption 2.6, the Euler characteristic equals the count of up-crossings of X about u and checking if X is above u at the start.

Definition 2.8 (**Euler Characteristic of Excursion Set**)**.** The Euler characteristic of the excursion set (2.2) is defined by

$$\varphi_{u,X}(0) = \mathbb{1}_{X(0) \geq u(0)} + N_{u,X}[0,1]. \tag{2.3}$$

We formulate the Kac-Rice formula in terms of the expected Euler characteristic $\mathbb{E}[\varphi_{u,X}(0)]$, that is, the expected number of times the process X is located above the function u in the sense of Definition 2.8 since that is our object of interest for this work. For the slightly more general result of Liebl and Reimherr (2023) where the correction location is an arbitrary location $t_0 \in [0,1]$, refer to Theorem 2.17. The simulation study with which we showed that t_0 has, in

fact, no influence on the outcome, at least in the case of linear critical value functions $u(t)$, can be found in the Appendix B.

The following theorem originates from Theorem 3.1, Liebl and Reimherr (2023), but was complemented with Assumption 2.6 and altered towards a fixed correction term $\mathbb{P}(X(0) \geq u(0))$.

Theorem 2.9 (**Generalized Kac-Rice Formula**). Let $X = \{X(t), t \in [0,1]\}$ be a centered elliptical stochastic process with $X \in C^1[0,1]$ almost surely as in Assumption 2.1, $u \in C^1_{a.e.}[0,1]$ as in Assumption 2.4 and let them together satisfy Assumption 2.6. Further, let $V > 0$ be the mixing coefficient of X such that $X(t) \overset{d}{=} VZ(t)$ (Lemma 2.3) for $Z(t) \sim \mathcal{N}(0,1)$, $t \in [0,1]$. Define $\mathcal{V} = V^{-2}$ along with its moment generating function, i.e. $M_{\mathcal{V}}(t) = \mathbb{E}[e^{t\mathcal{V}}]$. Consider the roughness parameter $\tau(t)$ that is defined by $\tau(t)^2 = \partial_{12} c(t,t) = \mathrm{Var}(Z'(t))$[2], where $c(t,s)$ is the dispersion function of $X(t)$ or equivalently the covariance function of $Z(t)$. Assume that $X(t)$ has a constant pointwise unit dispersion, i.e. $c(t,t) = 1$, and $\tau(t) > 0$ for all $t \in [0,1]$. Then, we have

$$\begin{aligned}\mathbb{E}[\varphi_{u,X}(0)] = &\mathbb{P}(X(0) \geq u(0)) + \int_0^1 \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\bigg(-\frac{1}{2}\bigg[u(t)^2 + \frac{u'(t)^2}{\tau(t)^2}\bigg]\bigg)\, dt \\ &- \int_0^1 \int_0^\infty \frac{u'(t)}{2\pi\tau(t)} M'_{\mathcal{V}}\bigg(-\frac{1}{2}\bigg[u(t)^2 + \frac{(y+u'(t))^2}{\tau(t)^2}\bigg]\bigg)\, dy\, dt.\end{aligned} \tag{2.4}$$

Proof. The proof primarily proceeds under the assumption that X is a Gaussian process, with the findings being extended to encompass elliptical processes in the final step.
We outline the main steps of the proof here and come back to them in Subsection 2.4 where we prove the theorem in complete detail.

1. Smooth approximation of up-crossing count: We show an integral representation of $N_{u,X}[0,1]$ using a smooth kernel K (counting formula, see Lemma 2.18). The fact that the up-crossing locations are isolated is important here.
2. Interpolation step for process $X(t)$ and critical value function $u(t)$: We build a linear interpolation $\tilde{X}_k(t)$ and $\tilde{u}_k(t)$ on a dyadic grid and prove that the counting formula of the previous step does also hold for $N_{\tilde{u}_k,\tilde{X}_k}[0,1]$ and is bounded from above (Lemma 2.19).
3. Expectation step: We take the expected value of $N_{\tilde{u}_k,\tilde{X}_k}[0,1]$ and obtain the expected number of up-crossings in the interpolation case (Equation (2.11)).
4. Refinement of the grid: We let $k \to \infty$ and pass the limit inside the integral. Therefore, we need to show that integrand of $\mathbb{E}[N_{\tilde{u}_k,\tilde{X}_k}[0,1]]$ is bounded. Plugging in the densities of the normal distribution yields the Kac-Rice formula for non-constant critical value functions and Gaussian processes (Lemma 2.20).
5. Generalization step for elliptical processes: Utilizing the characterization $X \overset{d}{=} VZ$ as per Lemma 2.3, we take advantage of the fact that the distribution of X conditioned on V is Gaussian. Thus, we again take the expectation over the mixing parameter V,

$$\mathbb{E}[\varphi_{u,X}(0)] = \mathbb{E}[\mathbb{E}[\varphi_{u,X}(0)|V]] = \mathbb{E}[\mathbb{E}[\varphi_{u/V,Z}(0)|V]],$$

to pass our result from Gaussian processes to elliptical processes.

□

[2] $\partial_{12} c(t,t)$ denotes the second partial derivative of c after the first and second argument evaluated in $(t,t) \in [0,1]^2$.

2.2. Special Cases

In this subsection, we provide the most relevant special cases of Theorem 2.9.

Corollary 2.10 (**Constant Critical Value Function**)**.** We are in the setting of Theorem 2.9 but we let the critical value function be constant, $u(t) \equiv u$. Then, the Kac-Rice formula (2.4) for an elliptical process $X(t)$ yields

$$\mathbb{E}[\varphi_{u,X}(0)] = \mathbb{P}(X(0) \geq u) + \frac{\|\tau\|_{L^1}}{2\pi} M_{\mathcal{V}}\Big(-\frac{u^2}{2}\Big). \tag{2.5}$$

Proof. If $u(t)$ is constant then $u'(t) = 0$, so the last two terms of Formula (2.4) vanish. Further, the moment generating function does not depend on t anymore and consequently, we can pull it out from the integral leaving $\int_0^1 \tau(t)dt = \|\tau\|_{L^1}$ since $\tau(t) \geq 0$ by assumption. Lastly, we use that $X(t) \overset{d}{=} X(0)$ in order to obtain the correction term $\mathbb{P}(X(0) \geq u)$. □

Corollary 2.11 (**Linear Critical Value Function**)**.** We remain in the setting of Theorem 2.9 but let the critical value function be linear, i.e. $u(t) = u(0) + at$, where a is the slope. Then, the Kac-Rice formula (2.4) for an elliptical process $X(t)$ yields

$$\begin{aligned}\mathbb{E}[\varphi_{u,X}(0)] =& \mathbb{P}(X(0) \geq u(0)) + \int_0^1 \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\Big(-\frac{1}{2}\Big[(u(0)+at)^2 + \frac{a^2}{\tau(t)^2}\Big]\Big)\, dt \\ &- \int_0^1 \int_0^\infty \frac{a}{2\pi\tau(t)} M'_{\mathcal{V}}\Big(-\frac{1}{2}\Big[(u(0)+at)^2 + \frac{(y+a)^2}{\tau(t)^2}\Big]\Big)\, dy\, dt.\end{aligned} \tag{2.6}$$

Liebl and Reimherr (2023) provide in their Corollaries 3.2 and 3.3 the Kac-Rice formula for more specific distribution assumption on the process $X(t)$ that we replicate in the following, but with regard to our Theorem 2.9, and we add the formula for linear critical value functions respectively.

Corollary 2.12 (**Gaussian Processes**)**.** In the setting of Theorem 2.9, we assume the process $X(t)$ to be mean-zero Gaussian and let $\mathrm{Var}(X(t)) = \sigma$. Then, the Kac-Rice formula (2.4) with adaptive critical value function $u(t)$ yields

$$\begin{aligned}\mathbb{E}[\varphi_{u,X}(0)] =& \Phi\Big(\frac{-u(0)}{\sigma}\Big) + \int_0^1 \frac{\tau(t)}{2\pi} \exp\Big(-\frac{1}{2\sigma^2}\Big[u(t)^2 + \frac{u'(t)^2}{\tau(t)^2}\Big]\Big)\, dt \\ &- \int_0^1 \frac{u'(t)}{\sqrt{2\pi\sigma^2}} \exp\Big(-\frac{u(t)^2}{2\sigma^2}\Big)\Phi\Big(\frac{-u'(t)}{\sigma\tau(t)}\Big)\, dt,\end{aligned}$$

where Φ is the cumulative distribution function of the standard normal. For a linear critical value function, $u(t) = u(0) + at$, we obtain

$$\begin{aligned}\mathbb{E}[\varphi_{u,X}(0)] =& \Phi\Big(\frac{-u(0)}{\sigma}\Big) + \int_0^1 \frac{\tau(t)}{2\pi} \exp\Big(-\frac{1}{2\sigma^2}\Big[(u(0)+at)^2 + \frac{a^2}{\tau(t)^2}\Big]\Big)\, dt \\ &- \int_0^1 \frac{a}{\sqrt{2\pi\sigma^2}} \exp\Big(-\frac{(u(0)+at)^2}{2\sigma^2}\Big)\Phi\Big(\frac{-a}{\sigma\tau(t)}\Big)\, dt.\end{aligned}$$

For a constant critical value function, $u(t) \equiv u$, we obtain the classic Kac-Rice formula

$$\mathbb{E}[\varphi_{u,X}(0)] = \Phi\Big(\frac{-u}{\sigma}\Big) + \frac{\|\tau\|_{L^1}}{2\pi} \exp\Big(-\frac{u^2}{2\sigma^2}\Big).$$

Proof. Apply the result of Lemma 2.20 to the critical value function $u(t)/\sigma$ and the process $X(t)/\sigma$ which consequently is standard Gaussian. The formula for a constant critical value functions follows similar to Corollary 2.10. □

Corollary 2.13 (*t*-**Processes**)**.** In the setting of Theorem 2.9, we assume the process X to be a t-process with ν degrees of freedom such that $\mathcal{V} \sim \chi^2_\nu/\nu$. Then, the Kac-Rice formula (2.4) with adaptive critical value function $u(t)$ yields

$$\begin{aligned}\mathbb{E}[\varphi_{u,X}(0)] = F_{t_\nu}(-u(0)) &+ \int_0^1 \frac{\tau(t)}{2\pi}\left(1 + \frac{u(t)^2}{\nu} + \frac{u'(t)^2}{\tau(t)^2\nu}\right)^{-\nu/2} dt \\ &- \int_0^1 \frac{u'(t)}{2\pi\tau(t)}\left(1 + \frac{u(t)^2}{\nu}\right)^{-\nu/2-1} \frac{\Gamma((\nu-1)/2)\sqrt{(\nu+1)\pi}a(t)}{\Gamma((\nu+2)/2)} F_{t_{\nu+1}}\left(\frac{-u'(t)}{a(t)}\right) dt,\end{aligned}$$

where F_{t_ν} is the cumulative distribution function of the t-distribution with ν degrees of freedom, Γ is the gamma function and $a(t)^2 := \nu\tau(t)^2(1 + u(t)^2/\nu)/(\nu+1)$.
For a constant critical value function, $u(t) \equiv u$, we obtain

$$\mathbb{E}[\varphi_{u,X}(0)] = F_{t_\nu}(-u) + \frac{\|\tau\|_{L^1}}{2\pi}\left(1 + \frac{u^2}{\nu}\right)^{-\nu/2}.$$

Proof. Refer to Liebl and Reimherr (2023), Section A.3. □

Remark 2.14. At the beginning of subsection 2.1, we restricted ourselves to the domain $[0,1]$ to calculate the Euler characteristic of the excursion set. But since, later on, we will apply the Kac-Rice formula over smaller intervals $[a_1, a_2] \subseteq [0,1]$, for completeness reasons, we want to remark how the formula looks like in this case. We stay in the setting of Theorem 2.9 and consider

$$\varphi_{u,X,[a_1,a_2]}(a_1) = \mathbb{P}(X(a_1) \geq u(a_1)) + N_{u,X}[a_1, a_2].$$

Then, analogously to Equation (2.4), we obtain

$$\begin{aligned}\mathbb{E}[\varphi_{u,X,[a_1,a_2]}(a_1)] = \mathbb{P}(X(a_1) \geq u(a_1)) &+ \int_{a_1}^{a_2} \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\left(-\frac{1}{2}\left[u(t)^2 + \frac{u'(t)^2}{\tau(t)^2}\right]\right) dt \\ &- \int_{a_1}^{a_2}\int_0^\infty \frac{u'(t)}{2\pi\tau(t)} M'_{\mathcal{V}}\left(-\frac{1}{2}\left[u(t)^2 + \frac{(y+u'(t))^2}{\tau(t)^2}\right]\right) dy\, dt.\end{aligned}$$

2.3. Generalization for Arbitrary Correction Location

We briefly introduce the version of the Kac-Rice formula that was originally proposed by Liebl and Reimherr (2023). It is slightly more general in the way that the location where we check whether X is located above u is not at the left interval boundary $t_0 = 0$ but at any arbitrary but fixed location $t_0 \in [0,1]$. This location is called correction location in the following.

This Kac-Rice formula with arbitrary t_0 will be relevant in the later discussions on constructing confidence bands, see Subsection 3.1, and in particular on the algorithms for computing fair confidence bands, as detailed in Subsection 5.2. Thus, we introduce it for sake of completeness. However, we will demonstrate that the correction location t_0 has, in fact, no influence on the results, at least in the case of constant and linear critical value functions. Further information can be found in Appendix B. Consequently, we will see that the Kac-Rice formula from Theorem 2.9 suffices for our purposes.

First of all, we define down-crossings equivalently to up-crossings in Definition 2.7.

Definition 2.15 (**Down-Crossings**)**.** The number of down-crossings of $X(t)$ about $u(t)$ on the interval $[a,b]$, $0 \leq a < b \leq 1$ is defined by

$$N^{-1}_{u,X}[a,b] = \#\{t \in [a,b] : X(t) = u(t),\ X'(t) < u'(t)\}.$$

The decisive point now is employing a dependence of the Euler characteristic on the correction location t_0.

Definition 2.16 (General Euler Characteristic of Excursion Set). The general Euler characteristic of excursion set (2.2) is

$$\varphi_{u,X}(t_0) = \mathbb{1}_{X(t_0)\geq u(t_0)} + N^{-1}_{u,X}[0,t_0] + N_{u,X}[t_0,1],$$

with some arbitrary mid point $t_0 \in [0,1]$.

The following theorem delivers a formula for $\mathbb{E}[\varphi_{u,X}(t_0)]$ in place of $\mathbb{E}[\varphi_{u,X}(0)]$, but apart from that, it is equivalent to Theorem 2.9. It is the version of the Kac-Rice formula that was originally introduced by Liebl and Reimherr (2023) in Theorem 3.1. The respective special cases can be found in Appendix A.2.

Theorem 2.17 (Generalized Kac-Rice Formula with Arbitrary t_0). Let $X = \{X(t), t \in [0,1]\}$ be a centered elliptical stochastic process with $X \in C^1[0,1]$ almost surely as in Assumption 2.1, $u \in C^1_{a.e.}[0,1]$ and let them together satisfy Assumption 2.6. Further, let $V > 0$ be the mixing coefficient of X such that $X(t) \overset{d}{=} VZ(t)$ (Lemma 2.3) for $Z(t) \sim \mathcal{N}(0,1)$, $t \in [0,1]$. Define $\mathcal{V} = V^{-2}$ along with its moment generating function, i.e. $M_{\mathcal{V}}(t) = \mathbb{E}[e^{t\mathcal{V}}]$. Consider the roughness parameter $\tau(t)$ that is defined by $\tau(t)^2 = \partial_{12}c(t,t) = \mathrm{Var}(Z'(t))$[3], where $c(t,s)$ is the dispersion function of $X(t)$ or equivalently the covariance function of $Z(t)$. Assume that $X(t)$ has a constant pointwise dispersion, i.e. $c(t,t) = 1$, and $\tau(t) > 0$ for all $t \in [0,1]$. Then, for any fixed $t_0 \in [0,1]$, we have

$$\begin{aligned}\mathbb{E}[\varphi_{u,X}(t_0)] =& \mathbb{P}(X(t_0) \geq u(t_0)) + \int_0^1 \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\bigg(-\frac{1}{2}\bigg[u(t)^2 + \frac{u'(t)^2}{\tau(t)^2}\bigg]\bigg)\,dt \\ &+ \int_0^{t_0}\int_0^\infty \frac{u'(t)}{2\pi\tau(t)} M'_{\mathcal{V}}\bigg(-\frac{1}{2}\bigg[u(t)^2 + \frac{(y-u'(t))^2}{\tau(t)^2}\bigg]\bigg)\,dy\,dt \\ &- \int_{t_0}^1\int_0^\infty \frac{u'(t)}{2\pi\tau(t)} M'_{\mathcal{V}}\bigg(-\frac{1}{2}\bigg[u(t)^2 + \frac{(y+u'(t))^2}{\tau(t)^2}\bigg]\bigg)\,dy\,dt.\end{aligned} \tag{2.7}$$

Proof. We generalize Theorem 2.9 for an arbitrary correction location $t_0 \in [0,1]$. Hence, consider the Euler characteristic as in Definition 2.16, that is

$$\varphi_{u,X}(t_0) = \mathbb{1}_{X(t_0)\geq u(t_0)} + N^{-1}_{u,X}[0,t_0] + N_{u,X}[t_0,1],$$

and notice that, when starting at some mid point t_0, we count crossings away from t_0 in both directions. Thus, when moving from t_0 towards 0, up-crossings as t decreases are equivalent to down-crossings as t increases. Also observe that the expected number of down-crossings about $u(t)$ equals the expected number of up-crossings about $-u(t)$ since we assume that X is centered which, along with the assumption of being elliptical, leads to this symmetry.
Therefore, we apply the theory of Theorem 2.9 for the expected number of up-crossing once to the interval $[t_0,1]$ with critical value function $u(t)$,

$$\begin{aligned}\mathbb{E}[N_{u,X}[t_0,1]] =& \int_{t_0}^1 \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\bigg(-\bigg[\frac{u(t)^2}{2} + \frac{u'(t)^2}{2\tau(t)^2}\bigg]\bigg)\,dt \\ &- \int_{t_0}^1\int_0^\infty \frac{u'(t)}{2\pi\tau(t)^2} M'_{\mathcal{V}}\bigg(-\bigg[\frac{u(t)^2}{2} + \frac{(y+u'(t))^2}{2\tau(t)^2}\bigg]\bigg)\,dy\,dt,\end{aligned}$$

[3] $\partial_{12}c(t,t)$ denotes the second partial derivative of c after the first and second argument evaluated in $(t,t) \in [0,1]^2$.

and once again to the interval $[0, t_0]$ with critical value function $-u(t)$,

$$\mathbb{E}[N_{u,X}^{-1}[0,t_0]] = \mathbb{E}[N_{-u,X}[0,t_0]] = \int_0^{t_0} \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\Big(- \Big[\frac{(-u(t))^2}{2} + \frac{(-u'(t))^2}{2\tau(t)^2}\Big]\Big)\, dt \\ - \int_0^{t_0}\int_0^{\infty} \frac{-u'(t)}{2\pi\tau(t)^2} M'_{\mathcal{V}}\Big(- \Big[\frac{(-u(t))^2}{2} + \frac{(y-u'(t))^2}{2\tau(t)^2}\Big]\Big)\, dy\, dt.$$

Together, this yields the claim

$$\mathbb{E}[\varphi_{u,X}(t_0)] = \mathbb{P}(X(0) \geq u(0)) + \int_0^1 \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\Big(- \Big[\frac{u(t)^2}{2} + \frac{u'(t)^2}{2\tau(t)^2}\Big]\Big)\, dt \\ + \int_0^{t_0}\int_0^{\infty} \frac{u'(t)}{2\pi\tau(t)^2} M'_{\mathcal{V}}\Big(- \Big[\frac{u(t)^2}{2} + \frac{(y-u'(t))^2}{2\tau(t)^2}\Big]\Big)\, dy\, dt \\ - \int_{t_0}^1\int_0^{\infty} \frac{u'(t)}{2\pi\tau(t)^2} M'_{\mathcal{V}}\Big(- \Big[\frac{u(t)^2}{2} + \frac{(y+u'(t))^2}{2\tau(t)^2}\Big]\Big)\, dy\, dt.$$

□

2.4. Proof of Theorem 2.9

We present the proof of Theorem 2.9 following the steps outlined in Subsection 2.1. Additionally, supplementary material required for this proof is provided in Appendix A.1. This is an extended and more detailed version of the proof given by Liebl and Reimherr (2023).

1. Smooth Approximation of Up-Crossing Count

The following lemma corresponds to Lemma A.2 (a) from Liebl and Reimherr (2023).

Lemma 2.18 (**Counting Formula**).
Let $X(t)$ be a Gaussian random process with $\mathrm{Var}(X(t)) > 0$ for every $t \in [0,1]$ and let $u(t)$ be a deterministic function. If $X \in C^1[0,1]$ almost surely, $u \in C^1_{\text{a.e.}}[0,1]$, then the number of up-crossings is

$$\begin{aligned} N_{u,X}[0,1] &= \#\{t \in [0,1] : X(t) = u(t),\, X'(t) - u'(t) > 0\} \\ &= \lim_{h\to 0} \int_0^1 \frac{1}{h} K\Big(\frac{u(t) - X(t)}{h}\Big)(X'(t) - u'(t))\mathbb{1}_{X'(t)-u'(t)>0}\, dt \\ &< \infty \quad \text{almost surely,} \end{aligned}$$

where K is a continuous, symmetric kernel function with compact support $[-1,1]$.

Proof. We start the proof with the case of no up-crossings, $N_{u,X}[0,1] = 0$. That is, we have either no crossing, $X(t) > u(t)$ for all $t \in [0,1]$ or $X(t) < u(t)$ for all $t \in [0,1]$, or we have exactly one down-crossing and no up-crossing.
Let us start with the former. Since $X(t) - u(t)$ is continuous and $[0,1]$ is compact, we can find an $\epsilon > 0$ that is smaller than the "smallest distance" between the functions, in symbols $|X(t) - u(t)| > \epsilon$ for all $t \in [0,1]$. So for a sufficiently small $h \leq \epsilon$, we get $K((u(t) - X(t))/h) = 0$ for all $t \in [0,1]$ because the kernel has support $[-1,1]$. This yields $N_{u,X}[0,1] = 0$.
In the latter case of exactly one down-crossing, we denote $t_d \in [0,1]$ the down-crossing location where $X(t_d) = u(t_d)$ and $X'(t_d) < u'(t_d)$ ($X'(t_d) = u'(t_d)$ is excluded by Assumption 2.6). This point is random and with probability zero falls on a discontinuity point of u' which is continuous almost everywhere. Thus, we can find a small neighborhood around t_d such that $\mathbb{1}_{X'(t)-u'(t)>0} = 0$ for all $t \in (t_d - \delta, t_d + \delta)$, $\delta > 0$, with probability one. Outside of this interval, it is $K((u(t) - X(t))/h) = 0$ for a sufficiently small h. Together, this implies that $N_{u,X}[0,1] = 0$

almost surely.

Now, we consider $N_{u,X}[0,1] = m$ with $m > 0$ and $X_u := X - u$ counting the up-crossings of X_u above 0 instead of the up-crossings of X above u,

$$\begin{aligned} N_{u,X}[0,1] &= \#\{t \in [0,1] : X(t) = u(t), X'(t) > u'(t)\} \\ &= \#\{t \in [0,1] : X_u(t) = 0, X_u'(t) > 0\} \qquad = m. \end{aligned}$$

X_u' is also continuous almost everywhere such that the up-crossing locations do not fall on discontinuity points of X_u' almost surely and the quantity $N_{u,X}[0,1]$ is well-defined. Further, $X_u(0) \neq 0$ and $X_u(1) \neq 0$ almost surely and therefore, we can label the up-crossing locations by $0 < t_1 < \cdots < t_m < 1$. Note that the points are isolated due to Assumption 2.6.

Let $I_1, \ldots, I_m$ be pairwise disjoint, open intervals that contain the up-crossing locations, $t_j \in I_j$ for $j = 1, \ldots, m$, and such that, restricted to those intervals, the derivative X_u' is positive. This exists since we have $X_u'(t_j) > 0$ due to the definition of up-crossings and since X_u' is almost surely continuous at the up-crossing locations. Hence, X_u restricted to I_j is almost surely a diffeomorphism for each $j = 1, \ldots, m$ respectively. Further, for sufficiently small h, Lemma A.1 yields that

$$X_u^{-1}(-h, h) \subseteq \bigcup_{j=1}^{m} I_j. \tag{2.8}$$

Now, we consider one up-crossing on the interval I_j where $\mathbb{1}_{X_u'(t)>0} \equiv 1$ and perform a change of variables with $y = X_u(t)/h$ and $dy = X_u'(t)/h\, dt$. For sufficiently small h, this leads to[4]

$$\int_{I_j} K\left(\frac{X_u(t)}{h}\right) \frac{X_u'(t)}{h}\, dt = \int_{-1}^{1} K(y)\, dy = 1 \quad \text{for every } j = 1, \ldots, m,$$

because K is a kernel function with support $[-1, 1]$. Note that the right hand side does not depend on h anymore and that the integral equals zero outside of $\cup_j I_j$ due to Equation (2.8). So putting the results together to the range of integration $[0, 1]$ leads to

$$\begin{aligned} N_{u,X}[0,1] = m &= \lim_{h \to 0} \sum_{j=1}^{m} \int_{I_j} K\left(\frac{X_u(t)}{h}\right) \frac{X_u'(t)}{h}\, dt \\ &= \lim_{h \to 0} \int_0^1 \frac{1}{h} K\left(\frac{X_u(t)}{h}\right) X_u'(t) \mathbb{1}_{X_u'(t)>0}\, dt. \end{aligned}$$

Finally, the number of up-crossings is almost surely finite because we assumed $X \in C^1[0,1]$ almost surely. □

2. Interpolation Step for Process $X(t)$ and Critical Value Function $u(t)$

A similar formula for counting the up-crossings can be found if we interpolate the process $X(t)$ and the critical value function $u(t)$ on a dyadic grid $\{j/2^k, j = 1, \ldots, 2^k\}$ for an arbitrary but

[4] compare to Adler and Taylor (2007), page 264

fixed integer k. For $t \in [0,1]$ such that $j-1 < t2^k \leq j$ where $j = 1, \dots, 2^k$ is an index, we let

$$\begin{aligned}
\tilde{X}_k(t) &= X\left(\frac{j}{2^k}\right)\left(2^k t - (j-1)\right) + X\left(\frac{j-1}{2^k}\right)\left(j - 2^k t\right),\\
\tilde{u}_k(t) &= u\left(\frac{j}{2^k}\right)\left(2^k t - (j-1)\right) + u\left(\frac{j-1}{2^k}\right)\left(j - 2^k t\right),\\
\tilde{X}_k'(t) &= 2^k\left(X\left(\frac{j}{2^k}\right) - X\left(\frac{j-1}{2^k}\right)\right),\\
\tilde{u}_k'(t) &= 2^k\left(u\left(\frac{j}{2^k}\right) - u\left(\frac{j-1}{2^k}\right)\right).
\end{aligned}$$

The following lemma corresponds to Lemma A.2 (b) from Liebl and Reimherr (2023).

Lemma 2.19 (**Counting Formula for Interpolation**).
Let all the requirements of Lemma 2.18 hold. Then the dyadic linear interpolations $\tilde{X}_k(t)$ and $\tilde{u}_k(t)$ satisfy

$$\begin{aligned}
N_{\tilde{u}_k,\tilde{X}_k}[0,1] &= \#\{t \in [0,1] : \tilde{X}_k(t) = \tilde{u}_k(t),\, \tilde{X}_k'(t) > \tilde{u}_k'(t)\}\\
&= \lim_{h\to 0} \int_0^1 \frac{1}{h} K\left(\frac{\tilde{u}_k(t) - \tilde{X}_k(t)}{h}\right)(\tilde{X}_k'(t) - \tilde{u}_k'(t))\mathbb{1}_{\tilde{X}_k'(t)-\tilde{u}_k'(t)>0}\, dt\\
&\leq \min\{2^k, N_{u,X}[0,1]\} \quad \text{almost surely.}
\end{aligned}$$

Proof. The main idea of this proof is the application of the counting formula of Lemma 2.18 to the intervals formed by the dyadic grid. Then, the two upper bounds follow naturally with the dyadic decomposition and with the intermediate value theorem.

First, we make sure that the up-crossings are not located on the grid points simply by noting that $X(t)$ is assumed to be Gaussian such that $X(j/2^k) \neq u(j/2^k)$ with probability one for each $j = 1, \dots, 2^k$. Hence, the up-crossing locations are almost surely within the partition intervals $((j-1)/2^k, j/2^k)$, $j = 1, \dots, 2^k$.
We remark that on each of the intervals, the process $\tilde{X}_k$ is just a very simple Gaussian process (namely the linear function through two centered Gaussian random variables) and $\tilde{u}_k$ is just a linear threshold. Thus, the requirements of Lemma 2.18 are fulfilled on each subinterval and we can apply Lemma 2.18 with $[(j-1)/2^k, j/2^k]$ instead of $[0,1]$. This yields

$$\begin{aligned}
N_{\tilde{u}_k,\tilde{X}_k}[0,1] &= \sum_{j=1}^{2^k} \#\left\{t \in \left[\frac{j-1}{2^k}, \frac{j}{2^k}\right] : \tilde{X}_k(t) = \tilde{u}_k(t),\, \tilde{X}_k'(t) > \tilde{u}_k'(t)\right\}\\
&= \sum_{j=1}^{2^k} \lim_{h\to 0} \int_{(j-1)/2^k}^{j/2^k} \frac{1}{h} K\left(\frac{\tilde{u}_k(t) - \tilde{X}_k(t)}{h}\right)(\tilde{X}_k'(t) - \tilde{u}_k'(t))\mathbb{1}_{\tilde{X}_k'(t)-\tilde{u}_k'>0}\, dt
\end{aligned}$$

almost surely such that the claim follows by exchanging the limit and the finite sum.

It remains to show the upper bounds. The functions $\tilde{X}_k$ and $\tilde{u}_k$ are linear within each subinterval and so is their difference $\tilde{X}_k - \tilde{u}_k$ that does, almost surely, not equal zero. Due to the fundamental theorem of algebra, there is at most one up-crossing per interval and since we have 2^k intervals, the upper bound 2^k for $N_{\tilde{u}_k,\tilde{X}_k}[0,1]$ comes naturally.
On the other hand, the number of up-crossings is bounded by the true number of up-crossings $N_{u,X}[0,1]$ because on each subinterval, we can at most miss up-crossings due to the linear. To formalize this, we consider an interval $[(j-1)/2^k, j/2^k]$ and imagine

$$(X-u)\left(\frac{j-1}{2^k}\right) < 0 \text{ and } (X-u)\left(\frac{j}{2^k}\right) > 0, \tag{2.9}$$

which leads to an up-crossing of $\tilde{X}_k$ above $\tilde{u}_k$. Keeping in mind that $X-u \in \mathcal{C}[0,1]$, the intermediate value theorem yields the existence of some $c \in [(j-1)/2^k, j/2^k]$ such that $(X-u)(c) = 0$. Due to the inequalities in (2.9), we can assume that we found an up-crossing of X about u in the interval. This up-crossing does not necessarily need to have location c and there must not necessarily be only one up-crossing in the considered interval. In all other constellations of the inequalities in (2.9), we count zero up-crossings of the linear interpolation on the subinterval but we can not draw any conclusion for $N_{u,X}$. This shows the inequality statement. □

3. Expectation Step

By taking the expected value of the counting formula in Lemma 2.19 and letting $k \to \infty$ in order to eliminate the interpolation, a Kac-Rice formula can already be derived, equivalently to Lemma A.3 in Liebl and Reimherr (2023). It already takes into account non-constant critical value functions but only holds Gaussian stochastic processes at this stage.

Lemma 2.20 (**Kac-Rice Formula for Non-Constant Critical Value Functions**). Let $X(t)$ be a continuously differentiable, mean-zero Gaussian process with covariance function $C(t,s)$ and constant pointwise unit variance, i.e. $\mathrm{Var}(X(t)) = 1$ for all $t \in [0,1]$. Let $u(t)$ be a continuously differentiable function on $[0,1]$ and define the roughness parameter function $\tau(t)$ by $\tau(t)^2 = \mathrm{Var}(X'(t)) = \partial_{12}C(t,t)$. Then, the expected value of the Euler characteristic $\varphi_{u,X}(0) = \mathbb{1}_{X(0)\geq u(0)} + N_{u,X}[0,1]$ is

$$\begin{aligned}\mathbb{E}[\varphi_{u,X}(0)] =& \mathbb{P}(X(0) \geq u(0)) + \int_0^1 \frac{\tau(t)}{2\pi} \exp\left(-\frac{1}{2}\left[u(t)^2 + \frac{u'(t)^2}{\tau(t)^2}\right]\right) dt \\ &- \int_0^1 \frac{u'(t)}{\sqrt{2\pi}} \exp\left(-\frac{u(t)^2}{2}\right)\Phi\left(\frac{-u'(t)}{\tau(t)}\right) dt,\end{aligned} \tag{2.10}$$

where Φ is the cumulative distribution function of the standard normal distribution.

The setting of the above lemma does already fulfill the requirements of Theorem 2.9 for the case of $V = 0$ such that $X(t) \overset{d}{=} Z(t)$.

Proof. This proof is structured as follows: As a first step, we will take the expectation of the Euler characteristic in the interpolation case with an arbitrary but fixed k. After deriving a suitable form for $\mathbb{E}[N_{\tilde{u}_k,\tilde{X}_k}[0,1]]$, we get rid of the interpolation and let $k \to \infty$. In doing so, we need to justify extensively why we can pass the limit under the two integrals. Finally, we will use the results of Theorem A.2 about the joint distribution of a Gaussian process X and its derivative X' in order to plug in the explicit form of the density function and deduce (2.10).

So using the linear interpolation $\tilde{X}_k$ and $\tilde{u}_k$ known from Lemma 2.19, taking the expected value of $\varphi_{\tilde{u}_k,\tilde{X}_k}(0) = \mathbb{1}_{\tilde{X}_k(0)\geq\tilde{u}_k(0)} + N_{\tilde{u}_k,\tilde{X}_k}[0,1]$ and applying $\tilde{u}_k(0) = u(0)$, $\tilde{X}_k(0) = X(0)$ almost surely leads to

$$\mathbb{E}[\varphi_{\tilde{u}_k,\tilde{X}_k}(0)] = \mathbb{P}(X(0) \geq u(0)) + \mathbb{E}[N_{\tilde{u}_k,\tilde{X}_k}[0,1]].$$

For the last term, $\mathbb{E}[N_{\tilde{u}_k,\tilde{X}_k}[0,1]]$, we plug in the counting formula representation from Lemma 2.19 which yields

$$\mathbb{E}[N_{\tilde{u}_k,\tilde{X}_k}[0,1]] = \mathbb{E}\left[\lim_{h\to 0}\int_0^1 \frac{1}{h}K\left(\frac{\tilde{u}_k(t)-\tilde{X}_k(t)}{h}\right)(\tilde{X}_k'(t)-\tilde{u}_k'(t))\mathbb{1}_{\tilde{X}_k'(t)-\tilde{u}_k'(t)>0}\,dt\right].$$

Taking the expectation over X corresponds to integrating out both $\tilde{X}_k$ and $\tilde{X}_k'$, for which $g_{\tilde{X}_k\tilde{X}_k'}$ denotes the joint density of $\tilde{X}_k$ and $\tilde{X}_k'$. Further, since having the uniform upper bound 2^k,

the dominated convergence theorem allows us to exchange the integral and the limit. Together with Fubini's theorem, this yields

$$\begin{aligned}\mathbb{E}[N_{\tilde{u}_k,\tilde{X}_k}[0,1]] &= \mathbb{E}\left[\lim_{h\to 0}\int_0^1 \frac{1}{h}K\left(\frac{\tilde{u}_k(t)-\tilde{X}_k(t)}{h}\right)(\tilde{X}_k'(t)-\tilde{u}_k'(t))\mathbb{1}_{\tilde{X}_k'(t)-\tilde{u}_k'(t)>0}\,dt\right]\\ &= \lim_{h\to 0}\int_0^1\int_{-\infty}^{\infty}\int_{\tilde{u}_k'(t)}^{\infty}\frac{1}{h}K\left(\frac{\tilde{u}_k(t)-x}{h}\right)(y-\tilde{u}_k'(t))g_{\tilde{X}_k(t)\tilde{X}_k'(t)}(x,y)\,dy\,dx\,dt.\end{aligned}$$

Substituting $x = zh + \tilde{u}_k$ yields, by symmetry and compactness of the kernel K,

$$\mathbb{E}[N_{\tilde{u}_k,\tilde{X}_k}[0,1]] = \lim_{h\to 0}\int_0^1\int_{-1}^1\int_{\tilde{u}_k'(t)}^{\infty}K(z)(y-\tilde{u}_k'(t))g_{\tilde{X}_k(t)\tilde{X}_k'(t)}(zh+\tilde{u}_k(t),y)\,dy\,dz\,dt.$$

As mentioned in Theorem A.2, the density $g_{\tilde{X}_k(t)\tilde{X}'_k(t)}$ is bivariate Gaussian for all $t \in [0,1]$, thus uniformly bounded and decays exponentially for increasing argument y, specifically faster than the linear term $(y-\tilde{u}_k'(t))$, and the kernel K is uniformly bounded by 1. Hence, the entire expression $K(z)(y-\tilde{u}_k'(t))g_{\tilde{X}_k(t)\tilde{X}_k'(t)}(zh+\tilde{u}_k,y)$ must have a uniform upper bound. Therefore, we can again exchange limit and integral with the dominated convergence theorem such that

$$\begin{aligned}\mathbb{E}[N_{\tilde{u}_k,\tilde{X}_k}[0,1]] &= \int_0^1\int_{-1}^1\int_{\tilde{u}_k'(t)}^{\infty}K(z)(y-\tilde{u}_k'(t))g_{\tilde{X}_k(t)\tilde{X}_k'(t)}(\tilde{u}_k(t),y)\,dy\,dz\,dt\\ &= \int_0^1\int_{\tilde{u}_k'(t)}^{\infty}(y-\tilde{u}_k'(t))g_{\tilde{X}_k(t)\tilde{X}'_k(t)}(\tilde{u}_k(t),y)\,dy\,dt.\end{aligned} \tag{2.11}$$

The second equality follows from pulling out $\int_{-1}^1 K(z)dz$ equals 1.

4. Refinement of the Grid

In the following, we need to let the fineness of our partition for the interpolation go to zero by letting $k \to \infty$ in order to obtain the exact formula without interpolation.
Referring to the proof of the upper bound in Lemma 2.19, we see that $N_{\tilde{u}_k,\tilde{X}_k}[0,1]$ is monotonically increasing with k. Hence, due to the monotone convergence theorem, it is $\mathbb{E}[N_{u,X}[0,1]] = \lim_{k\to\infty}\mathbb{E}[N_{\tilde{u}_k,\tilde{X}_k}[0,1]]$. In order to let $k \to \infty$ in our integral expression (2.11), we rewrite it as

$$\mathbb{E}[N_{\tilde{u}_k,\tilde{X}_k}[0,1]] = \int_0^1 \mathbb{E}\big[(\tilde{X}_k'(t)-\tilde{u}_k'(t))\mathbb{1}_{\tilde{X}_k'(t)\geq\tilde{u}_k'(t)}|\tilde{X}_k(t)=\tilde{u}_k(t)\big]g_{\tilde{X}_k(t)}(\tilde{u}_k(t))\,dt. \tag{2.12}$$

We used here that $\tilde{X}_k'(t)$ and $\tilde{X}_k(t)$ are pointwise independent, as proven in Theorem A.2, so that their joint density decomposes into the product of the individual densities. In the following, we justify that this integrand is uniformly bounded in order to use the dominated convergence theorem again to pass $\lim_{k\to\infty}$ inside the integral.
We rewrite the linear interpolation $\tilde{X}_k(t)$ for $t \in [(j-1)/2^k, j/2^k]$ and $j = 1,\dots,2^k$ with non-random weights $w_t = (j-2^k t) \in [0,1]$ such that

$$\tilde{X}_k(t) = w_t X\left(\frac{j-1}{2^k}\right) + (1-w_t)X\left(\frac{j}{2^k}\right). \tag{2.13}$$

Since by assumption, the covariance of X is differentiable, thus in particular continuous for points $(t,t) \in [0,1]^2$ and $\mathrm{Cov}(X(t),X(t)) = \mathrm{Var}(X(t)) = 1$, there exists a k_0 such that for any $k \geq k_0$ it is $\mathrm{Cov}(X((j-1)/2^k), X(j/2^k)) \geq 0$. Thus

$$\begin{aligned}\mathrm{Var}(\tilde{X}_k(t)) &= w_t^2\mathrm{Var}\big(X\left(\frac{j-1}{2^k}\right)\big) + (1-w_t)^2\mathrm{Var}\big(X\left(\frac{j}{2^k}\right)\big)\\ &\quad + 2w_t(1-w_t)\mathrm{Cov}\big(X\left(\frac{j-1}{2^k}\right), X\left(\frac{j}{2^k}\right)\big)\\ &\geq w_t^2 + (1-w_t)^2 \geq 0.5.\end{aligned}$$

Consequently, for a sufficiently large k, the density $g_{\tilde{X}_k(t)}$ is uniformly bounded for all $t \in [0,1]$ by $(2\pi)^{-1/2}$ because

$$g_{\tilde{X}_k(t)}(y) = \frac{1}{\sqrt{2\pi}} \exp\Big(- \frac{y^2}{2\mathrm{Var}(\tilde{X}_k(t))^{1/2}} \Big) \leq \frac{1}{\sqrt{2\pi}} \exp\Big(- \frac{y^2}{2 \cdot 0.5^2} \Big) \leq \frac{1}{\sqrt{2\pi}}.$$

Further, by Theorem A.2, the joint distribution $(\tilde{X}_k(t), \tilde{X}_k'(t))$ is multivariate, mean-zero Gaussian and therefore, we can write

$$\tilde{X_k'}(t) = b_k(t)\tilde{X}_k(t) + \epsilon_k(t), \tag{2.14}$$

where $b_k(t) = \mathbb{E}[\tilde{X}_k(t)\tilde{X}_k'(t)]/\mathbb{E}[\tilde{X}_k(t)^2] = \mathrm{Cov}(\tilde{X}_k(t), \tilde{X}_k'(t))/\mathrm{Var}(\tilde{X}_k(t))$, and $\epsilon_k(t)$ and $\tilde{X}_k(t)$ are independent for each $t \in [0,1]$. With this, we can construct the bound

$$\begin{aligned}
&\mathbb{E}\big[(\tilde{X}_k'(t) - \tilde{u}_k'(t))\mathbb{1}_{\tilde{X}_k'(t) \geq \tilde{u}_k'(t)} | \tilde{X}_k(t) = \tilde{u}_k(t)\big] \\
&\leq \mathbb{E}\big[\, \Big| \tilde{X}_k'(t) - \tilde{u}_k'(t) \Big| \, | \tilde{X}_k(t) = \tilde{u}_k(t)\big] \\
&= \mathbb{E}\big[\, \Big| b_k(t)\tilde{X}_k(t) + \epsilon_k(t) - \tilde{u}_k'(t) \Big| \, | \tilde{X}_k(t) = \tilde{u}_k(t)\big] \\
&\leq \big| b_k(t)\tilde{u}_k(t) - \tilde{u}_k'(t) \big| + \mathbb{E}[\epsilon_k(t)].
\end{aligned} \tag{2.15}$$

We show separately that those two terms are bounded and start with the latter one.
Squaring and taking the expected value in (2.14) yields

$$\mathbb{E}[\tilde{X_k'}(t)^2] = b_k(t)^2 \mathbb{E}[\tilde{X}_k(t)^2] + \mathbb{E}[\epsilon_k(t)^2] \tag{2.16}$$

due to the independence of $\epsilon_k(t)$ and $\tilde{X}_k(t)$ and because $\tilde{X}_k(t)$ is centered. Hence

$$\mathbb{E}[\epsilon_k(t)^2] \leq \mathbb{E}[\tilde{X}_k'(t)^2] \quad \text{and} \quad \mathbb{E}[\epsilon_k(t)] \leq \mathbb{E}[\tilde{X}_k'(t)^2]^{\frac{1}{2}}$$

by Jensen's inequality. Remembering that $\tilde{X}_k'(t) = 2^k\big(X(j/2^k) - X((j-1)/2^k)\big)$, the mean value theorem delivers the existence of $\xi \in [(j-1)/2^k, j/2^k]$ such that $\tilde{X}_k'(t) = X'(\xi)$. Therefore,

$$\tilde{X}_k'(t)^2 \leq \big(\sup_{s \in [0,1]} X'(s) \big)^2, \quad t \in [0,1],$$

which has a finite expectation by Theorem 2.9 of Azaïs et al. (2002). This finally bounds $\mathbb{E}[\epsilon_k(t)^2]$ in (2.15).
In order to bound the other term in (2.15), we note that, by construction, $\tilde{u}_k$ and $\tilde{u}_k'$ have the following upper bound

$$\sup_{t \in [0,1]} |\tilde{u}_k(t)| \leq \sup_{t \in [0,1]} |u(t)| < \infty, \qquad \sup_{t \in [0,1]} |\tilde{u}_k'(t)| \leq \sup_{t \in [0,1]} |u'(t)| < \infty.$$

Furthermore, as argued before, $\mathrm{Var}(\tilde{X}_k(t)) \geq 0.5$ which implies $\mathbb{E}[\tilde{X}_k(t)^2] \geq 0.5$ which we plug in to the following inequality obtained from Equation (2.16) such that

$$b_k(t)^2 \leq \frac{\mathbb{E}\big[\tilde{X}_k'(t)^2\big]}{\mathbb{E}\big[\tilde{X}_k(t)^2\big]} \leq 2\mathbb{E}\big[\tilde{X}_k'(t)^2\big] \leq 2\mathbb{E}\big[\big(\sup_{s \in [0,1]} X'(s) \big)^2\big] < \infty.$$

Hence, the term $|b_k(t)\tilde{u}_k(t) - \tilde{u}_k'(t)|$ is uniformly bounded.
Overall, we showed that there exists a bound for the integrand in Equation (2.12) uniformly over all $t \in [0,1]$ for sufficiently large k so that we can exchange limit and integral by dominated convergence theorem. This yields

$$\mathbb{E}[N_{u,X}[0,1]] = \int_0^1 \lim_{k \to \infty} \Big[\mathbb{E}\big[(\tilde{X}_k'(t) - \tilde{u}_k'(t))\mathbb{1}_{\tilde{X}_k'(t) \geq \tilde{u}_k'(t)} | \tilde{X}_k(t) = \tilde{u}_k(t)\big] g_{\tilde{X}_k(t)}(\tilde{u}_k(t)) \Big] \, dt. \tag{2.17}$$

Now, having the limit in front of the integrand in (2.17), we aim to ensure the convergence of the integrand for all or almost all $t \in [0,1]$. We get the convergence of the density, that is,

$$g_{\tilde{X}_k(t)}(\tilde{u}_k(t)) \to \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{u(t)^2}{2} \Big) = g_{X(t)}(u(t)), \tag{2.18}$$

since $\tilde{u}_k(t) \to u(t)$ and $\mathrm{Var}(\tilde{X}_k(t)) \to \mathrm{Var}(X(t)) = 1$. This holds because, considering the representation in (2.13),

$$\mathrm{Cov}\big(X\Big(\frac{j-1}{2^k}\Big), X\Big(\frac{j}{2^k}\Big)\big) \to \mathrm{Cov}\big(X\Big(\frac{j}{2^k}\Big), X\Big(\frac{j}{2^k}\Big)\big) = \mathrm{Var}\big(X\Big(\frac{j}{2^k}\Big)\big) = 1$$
$$\text{and} \quad w_t^2 + (1-w_t)^2 + 2w_t(1-w_t) = 1.$$

For the conditional expected value term, we have similarly to Equation (2.15) that

$$\begin{aligned}
&\mathbb{E}\big[(\tilde{X}_k'(t) - \tilde{u}_k'(t))\mathbb{1}_{\tilde{X}_k'(t) \geq \tilde{u}_k'(t)} | \tilde{X}_k(t) = \tilde{u}_k(t)\big] \\
&= \mathbb{E}\big[(b_k(t)\tilde{u}_k(t) + \epsilon_k(t) - \tilde{u}_k'(t))\mathbb{1}_{b_k(t)\tilde{u}_k(t)+\epsilon_k(t) \geq \tilde{u}_k'(t)}\big] \\
&\leq \mathbb{E}\big[\,\big|b_k(t)\tilde{u}_k(t) + \epsilon_k(t) - \tilde{u}_k'(t)\big|\,\big].
\end{aligned}$$

Above, we have already seen that $b_k(t)$, $\tilde{u}_k(t)$ and $\tilde{u}_k'(t)$ are uniformly bounded for $k \geq k_0$ for a $k_0 \geq 1$ and over $t \in [0,1]$. Thus, again by the dominated convergence theorem, we can pass $\lim_{k\to\infty}$ under the integral of the expectation which yields

$$\begin{aligned}
&\lim_{k\to\infty} \mathbb{E}\big[(b_k(t)\tilde{u}_k(t) + \epsilon_k(t) - \tilde{u}_k'(t))\mathbb{1}_{b_k(t)\tilde{u}_k(t)+\epsilon_k(t) \geq \tilde{u}_k'(t)}\big] \\
&= \mathbb{E}\big[(b(t)u(t) + \epsilon(t) - u'(t))\mathbb{1}_{b(t)u(t)+\epsilon(t) \geq u'(t)}\big] \\
&= \mathbb{E}\big[(X'(t) - u'(t))\mathbb{1}_{X'(t) \geq u'(t)} | X(t) = u(t)\big].
\end{aligned}$$

Finally, plugging this and (2.18) back into Equation (2.17) results in

$$\begin{aligned}
\mathbb{E}[N_{u,X}[0,1]] &= \int_0^1 \mathbb{E}\big[(X'(t) - u'(t))\mathbb{1}_{X'(t) \geq u'(t)} | X(t) = u(t)\big] g_{X(t)}(u(t))\, dt \\
&= \int_0^1 \int_{u'(t)}^{\infty} (y - u'(t)) g_{X(t)X'(t)}(u(t), y)\, dy\, dt
\end{aligned}$$

where in the second equation, the conditional expectation was rewritten.

In a last step, we plug in the density function of the jointly mean-zero Gaussian random variable $(X(t), X'(t))$. Since we assumed $X(t)$ to have constant variance, by Theorem A.2, $X(t)$ and $X'(t)$ are pointwise uncorrelated and thus independent and $g_{X(t)X'(t)}(u(t), y) = g_{X(t)}(u(t)) g_{X'(t)}(y)$. Therefore, we obtain

$$\begin{aligned}
\mathbb{E}[\varphi_{u,X}(0)] &= \mathbb{P}(X(0) \geq u(0)) + \\
&\int_0^1 \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{u(t)^2}{2} \Big) \int_{u'(t)}^{\infty} (y - u'(t)) \frac{1}{\sqrt{2\pi\tau(t)^2}} \exp\Big(-\frac{y^2}{2\tau(t)^2} \Big)\, dy\, dt.
\end{aligned}$$

Here, $\tau(t) = \mathrm{Var}(X'(t))$ and the density $g_{X(t)}(u(t))$ got pulled out of the inner integral. To calculate the inner part, we split the integral, apply a change of variable $y = \tau(t)x$ and use

$\Phi(-x) = 1 - \Phi(x)$ such that

$$\begin{aligned}
&\int_{u'(t)}^{\infty} (y - u'(t)) \frac{1}{\sqrt{2\pi\tau(t)^2}} \exp\Big(- \frac{y^2}{2\tau(t)^2} \Big)\, dy \\
&= \int_{u'(t)}^{\infty} \frac{y}{\sqrt{2\pi\tau(t)^2}} \exp\Big(- \frac{y^2}{2\tau(t)^2} \Big)\, dy - u'(t) \int_{u'(t)}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\Big(- \frac{y^2}{2\tau(t)^2} \Big) \frac{1}{\tau(t)}\, dy \\
&= \Big[- \frac{\tau(t)}{\sqrt{2\pi}} \exp\Big(- \frac{y^2}{2\tau(t)^2} \Big) \Big]_{y=u'(t)}^{y=\infty} - u'(t) \int_{u'(t)/\tau(t)}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\Big(- \frac{x^2}{2} \Big)\, dx \\
&= \frac{\tau(t)}{\sqrt{2\pi}} \exp\Big(- \frac{u'(t)^2}{2\tau(t)^2} \Big) - u'(t) \Phi\Big(- \frac{u'(t)}{\tau(t)} \Big).
\end{aligned}$$

This finally yields

$$\begin{aligned}
\mathbb{E}[\varphi_{u,X}(0)] = \mathbb{P}(X(0) \geq u(0)) &+ \int_0^1 \frac{\tau(t)}{2\pi} \exp\Big(- \frac{1}{2} \Big[u(t)^2 + \frac{u'(t)^2}{\tau(t)^2} \Big] \Big)\, dt \\
&- \int_0^1 \frac{u'(t)}{\sqrt{2\pi}} \exp\Big(- \frac{u(t)^2}{2} \Big) \Phi\Big(\frac{-u'(t)}{\tau(t)} \Big)\, dt.
\end{aligned}$$

□

5. Generalization Step for Elliptical Processes

Now, with Lemma 2.20, we have a result for $\mathbb{E}[\varphi_{u,X}(0)]$ for a variable critical value function $u(t)$ but only for a Gaussian stochastic process $X(t)$. In the following, we will trace the elliptical case back to the Gaussian case and use the former result in order to obtain a Kac-Rice formula for $\mathbb{E}[\varphi_{u,X}(0)]$ where $X(t)$ is elliptical.

Proof of Theorem 2.9. In order to continue the previous results and ultimately prove the theorem, we exploit the characterization of elliptical processes from Lemma 2.3 that an elliptical stochastic process X can be expressed as a scalar mixture of a Gaussian process Z. This means we have $X(t) \overset{d}{=} VZ(t)$ pointwise for a random variable $V > 0$. Since so far, we obtained results for Gaussian processes, we now imagine that they hold conditioned on the scalar mixture coefficient V which is a random variable. Then, in order to move on to elliptical processes, we just need to take the expectation again.

For convenience, we use the alternative parametrization $V = \mathcal{V}^{-\frac{1}{2}}$ such that $X \overset{d}{=} \mathcal{V}^{-\frac{1}{2}} Z$ or equivalently $X\mathcal{V}^{\frac{1}{2}} \overset{d}{=} Z$. Note that if we let $v(t) = \mathcal{V}^{\frac{1}{2}} u(t)$ then by Definitions 2.7 and 2.8, we have

$$\varphi_{u,X}(0) = \varphi_{u,\mathcal{V}^{-\frac{1}{2}} Z}(0) = \varphi_{v,Z}(0).$$

Hence, with Lemma 2.20, we obtain

$$\begin{aligned}
\mathbb{E}[\varphi_{u,X}(0)|\mathcal{V}] &= \mathbb{E}[\varphi_{v,X\mathcal{V}^{\frac{1}{2}}}(0)|\mathcal{V}] = \mathbb{E}[\varphi_{v,Z}(0)|\mathcal{V}] \\
&= \mathbb{P}(Z(0) \geq v(0)|\mathcal{V}) + \int_0^1 \frac{\tau(t)}{2\pi} \exp\Big(- \frac{1}{2} \Big[v(t)^2 + \frac{v'(t)^2}{\tau(t)^2} \Big] \Big)\, dt \\
&\quad - \int_0^1 \frac{v'(t)}{\sqrt{2\pi}} \exp\Big(- \frac{v(t)^2}{2} \Big) \Phi\Big(\frac{-v'(t)}{\tau(t)} \Big)\, dt \\
&= \mathbb{P}(X(0) \geq u(0)|\mathcal{V}) + \int_0^1 \frac{\tau(t)}{2\pi} \exp\Big(- \frac{\mathcal{V}}{2} \Big[u(t)^2 + \frac{u'(t)^2}{\tau(t)^2} \Big] \Big)\, dt \\
&\quad - \int_0^1 \frac{u'(t)\mathcal{V}^{\frac{1}{2}}}{\sqrt{2\pi}} \exp\Big(- \frac{\mathcal{V} u(t)^2}{2} \Big) \Phi\Big(\frac{-u'(t)\mathcal{V}^{\frac{1}{2}}}{\tau(t)} \Big)\, dt.
\end{aligned} \tag{2.19}$$

For the second equation, we inserted the definition of $v(t)$. Now, while taking the expected value over $\mathcal{V}$ yields $\mathbb{E}[\varphi_{u,X}(0)]$ on the left side and $\mathbb{P}(X(0) \geq u(0))$ for the first term on the right side. For the second term on the right, using Fubini's theorem term to interchange the integrals leads to

$$\int_0^1 \frac{\tau(t)}{2\pi} \int_0^\infty \exp\Big(- x\Big[\frac{u(t)^2}{2} + \frac{u'(t)^2}{2\tau(t)^2}\Big]\Big)\, dF_{\mathcal{V}}(x)\, dt,$$

where $F_{\mathcal{V}}$ is the cumulative distribution function of the random variable $\mathcal{V}$ that is strictly positive by definition. Notice that, therefore, we have a moment generating function (mgf) $M_{\mathcal{V}}(t) = \mathbb{E}[\exp(t\mathcal{V})]$ which will be finite since our argument is strictly negative and $\mathcal{V} > 0$. Hence for the second term, we obtain

$$\int_0^1 \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\Big(- \Big[\frac{u(t)^2}{2} + \frac{u'(t)^2}{2\tau(t)^2}\Big]\Big)\, dt.$$

Before taking the expected value of the third component, we dissolve the integral of the cumulative distribution function Φ by substituting $y = \tau(t)x/\mathcal{V}^{1/2} - u'(t)$ such that

$$\begin{aligned}
&- \int_0^1 \frac{u'(t)\mathcal{V}^{\frac{1}{2}}}{\sqrt{2\pi}} \exp\Big(- \frac{\mathcal{V}u(t)^2}{2}\Big)\Phi\Big(\frac{-u'(t)\mathcal{V}^{\frac{1}{2}}}{\tau(t)}\Big)\, dt \\
&= - \int_0^1 \frac{u'(t)\mathcal{V}^{\frac{1}{2}}}{\sqrt{2\pi}} \exp\Big(- \frac{\mathcal{V}u(t)^2}{2}\Big) \int_{\frac{u'(t)\mathcal{V}^{\frac{1}{2}}}{\tau(t)}}^\infty \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{x^2}{2}\Big)\, dx\, dt \\
&= - \int_0^1 \frac{u'(t)\mathcal{V}^{\frac{1}{2}}}{\sqrt{2\pi}} \exp\Big(- \frac{\mathcal{V}u(t)^2}{2}\Big) \int_0^\infty \frac{\mathcal{V}^{\frac{1}{2}}}{\sqrt{2\pi\tau(t)^2}} \exp\Big(- \frac{\mathcal{V}(y+u'(t))^2}{2\tau(t)^2}\Big)\, dy\, dt \\
&= - \int_0^1 \int_0^\infty \frac{u'(t)\mathcal{V}}{2\pi\tau(t)^2} \exp\Big(- \mathcal{V}\Big[\frac{u(t)^2}{2} + \frac{(y+u'(t))^2}{2\tau(t)^2}\Big]\Big)\, dy\, dt.
\end{aligned}$$

Then, we take the expected value over $\mathcal{V}$ and use Fubini's theorem to interchange the integrals

$$\begin{aligned}
&- \int_0^\infty \int_0^1 \int_0^\infty \frac{u'(t)x}{2\pi\tau(t)^2} \exp\Big(- x\Big[\frac{u(t)^2}{2} + \frac{(y+u'(t))^2}{2\tau(t)^2}\Big]\Big)\, dy\, dt\, dF_{\mathcal{V}}(x) \\
&\overset{\text{Fubini}}{=} - \int_0^1 \int_0^\infty \frac{u'(t)}{2\pi\tau(t)^2} \int_0^\infty x \exp\Big(- x\Big[\frac{u(t)^2}{2} + \frac{(y+u'(t))^2}{2\tau(t)^2}\Big]\Big)\, dF_{\mathcal{V}}(x)\, dy\, dt \\
&\overset{\text{mgf}}{=} - \int_0^1 \int_0^\infty \frac{u'(t)}{2\pi\tau(t)^2} M'_{\mathcal{V}}\Big(- \Big[\frac{u(t)^2}{2} + \frac{(y+u'(t))^2}{2\tau(t)^2}\Big]\Big)\, dy\, dt,
\end{aligned}$$

where $M'_{\mathcal{V}}(t) = \mathbb{E}[\mathcal{V}\exp(t\mathcal{V})]$ (Leibniz integral rule).

This way, by incorporating the results into Equation (2.19), we obtain the form for elliptical distributions

$$\begin{aligned}
\mathbb{E}[\varphi_{u,X}(0)] = \mathbb{P}(X(0) \geq u(0)) &+ \int_0^1 \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\Big(- \Big[\frac{u(t)^2}{2} + \frac{u'(t)^2}{2\tau(t)^2}\Big]\Big)\, dt \\
&- \int_0^1 \int_0^\infty \frac{u'(t)}{2\pi\tau(t)^2} M'_{\mathcal{V}}\Big(- \Big[\frac{u(t)^2}{2} + \frac{(y+u'(t))^2}{2\tau(t)^2}\Big]\Big)\, dy\, dt.
\end{aligned}$$

□

3. Confidence Bands for Functional Parameter

With the Kac-Rice formula developed in the previous section, simultaneous confidence bands for a functional parameter $\theta(t)$, $t \in [0,1]$, can be constructed. This section is dedicated to explaining such a procedure, particularly focusing on parameters with asymptotically normal estimators, to broaden the scope of its application. Additionally, it covers the mean function as a special case.

Subsection 3.1 delivers the fundamental logic of constructing confidence bands with the Kac-Rice formula. Subsection 3.3 justifies that we can still apply this approach even if our distribution assumption on the processes is only asymptotically fulfilled. This is especially motivated by the asymptotic normality of various functional parameter estimators which will be dealt with in Subsection 3.2. There, we will also briefly explore the implications of this asymptotic result on potential functional parameters suitable for our approach. However, we finally restrict ourselves to the case of confidence bands for the mean function. This case will be discussed more thoroughly in Subsection 3.4.

3.1. Constructing Simultaneous Confidence Bands Using Kac-Rice Formula

The main idea in order to construct a simultaneous confidence band, introduced by Liebl and Reimherr (2023), is to consider a standardized version $X(t)$ of the parameter function $\theta(t)$ and to derive a critical value function $u(t)$ by solving the Kac-Rice formula (2.4) for this standardized process $X(t)$ and the given level of $\alpha/2$. The resulting critical value function then serves as the boundaries of the band for $X(t)$ and inverting the standardization then yields the boundaries for the band for $\theta(t)$.

Recall again the difference between pointwise confidence intervals and simultaneous confidence bands. While a classic pointwise confidence interval tells us only about the uncertainty at a fixed location, simultaneous confidence bands provide functional upper and lower bounds such that the entire true parameter function $\theta(t)$ is covered with a high probability. Thus, a simultaneous $(1-\alpha)$-confidence band $[\theta_l, \theta_u]$ for the functional parameter θ fulfills

$$\mathbb{P}(\forall\, t \in [0,1] : \theta_l(t) \leq \theta(t) \leq \theta_u(t)) \geq 1-\alpha.$$

So in order to build a confidence band for an arbitrary functional parameter $\theta(t)$, we need a functional estimator $\hat{\theta}(t)$ that we center correctly with the true parameter function $\theta(t)$ and standardize with the deviation of the estimator. Thus, we obtain the process

$$X(t) = \frac{\hat{\theta}(t) - \theta(t)}{(\mathrm{Var}(\hat{\theta}(t)))^{1/2}}. \tag{3.1}$$

Under appropriate distribution assumptions on $\hat{\theta}$ or rather the observed sample, one can derive that $X(t)$ is an elliptical process such that Assumption 2.1 is satisfied. We will do this in Subsection 3.4 for $\theta(t)$ being the mean function. Alternatively, it also suffices for X to be asymptotically elliptical which we address specifically in Subsection 3.2 and 3.3.
For such a process $X(t)$, we need to find a critical value function $u(t)$ in the sense of Definition 2.5 that satisfies

$$\mathbb{P}\big(\exists\, t \in [0,1] : X(t) \geq u(t)\big) \leq \alpha/2. \tag{3.2}$$

The following inequality, introduced as Equation (1) by Liebl and Reimherr (2023), delivers a useful upper bound for the excursion probability (3.2).

Theorem 3.1 (**Expected Euler Characteristic Inequality**)**.** Let $X(t)$ be a stochastic process in $C^1[0,1]$ and $u(t)$ be a critical value function defined on $[0,1]$ and let Assumptions 2.1 and 2.6 be satisfied, then

$$\mathbb{P}\big(\exists\, t \in [0,1] : X(t) \geq u(t)\big) \leq \mathbb{E}\big[\varphi_{u,X}(0)\big] \tag{3.3}$$

where $\varphi_{u,X}(0)$ is the Euler characteristic as in Definition 2.8.

Proof. Due to Assumptions 2.1 and 2.6, the number of up-crossings is well-defined. As shown in more detail in the proof of Theorem D.1, it is $\{\exists\, t \in [0,1] : X(t) \geq u(t)\} = \{X(0) \geq u(0)\} \cup \{N_{u,X}[0,1] \geq 1\}\}$. Thus, Boole's inequality and Markov's inequality yield

$$\begin{aligned}
\mathbb{P}(\exists\, t \in [0,1] : X(t) \geq u(t)) = \quad & \mathbb{P}\big(\{X(0) \geq u(0)\} \cup \{N_{u,X}[0,1] \geq 1\}\big) \\
\overset{\text{Boole}}{\leq} \quad & \mathbb{P}\big(X(0) \geq u(0)\big) + \mathbb{P}\big(N_{u,X}[0,1] \geq 1\big) \\
\overset{\text{Markov}}{\leq} \quad & \mathbb{P}\big(X(0) \geq u(0)\big) + \mathbb{E}\big[N_{u,X}[0,1]\big] \\
= \quad & \mathbb{E}\big[\varphi_{u,X}(0)\big].
\end{aligned}$$

The last equation is the definition of the Euler characteristic (2.3). □

Indeed, the Kac-Rice formula is of considerable importance as it has long been used as an approximation of an excursion probability. This is due to the fact that, from a heuristic perspective, if a process X crosses a high level u, it is unlikely to do so more than once.[5]

By means of this inequality (3.3), we can find a suitable critical value function $u(t)$ for (3.2) by equating

$$\mathbb{E}[\varphi_{u,X}(0)] = \alpha/2 \tag{3.4}$$

and solving for u using the Kac-Rice formula from Theorem 2.9. This is the central step in our construction where we see that the Kac-Rice formula plays a crucial role. Unfortunately, there is no way of generally simplifying the formula towards an explicit formula for u meaning that, in order to practically obtain the function $u(t)$, we must solve the respective Kac-Rice formula, e.g. one of those presented in Subsection 2.2, with numerical methods.
Besides, the result of (3.4) is usually not unique but a whole set of possible critical value functions. Thus, we only allow critical value functions within a set $\mathcal{U}$ that is convex, compact and contains the constant functions (up to an appropriate limit). We denote this solution set for Equation (3.4) with

$$\mathcal{U}_{\alpha/2}([0,1]) = \{u \in \mathcal{U} : \mathbb{E}[\varphi_{u,X}(0)] = \alpha/2\} \tag{3.5}$$

for a given $\alpha \in [0,1]$.
Let a resulting critical value function be denoted by $u_{\alpha/2} \in \mathcal{U}_{\alpha/2}([0,1])$. Then, the inequality (3.2) holds also for the excursion of the process $-X$ above $u_{\alpha/2}$. This is because $-X$ is distributed just like X and, due to symmetry, we have

$$\mathbb{P}\big(\exists\, t \in [0,1] : -X(t) \geq u_{\alpha/2}(t)\big) \leq \mathbb{E}[\varphi_{u_{\alpha/2},-X}(0)] = \mathbb{E}[\varphi_{u_{\alpha/2},X}(0)] = \alpha/2.$$

Hence, together with (3.3), we can deduce that

$$\begin{aligned}
& \mathbb{P}\big(\exists\, t \in [0,1] : |X(t)| \geq |u_{\alpha/2}(t)|\big) \\
= \; & \mathbb{P}\big(\exists\, t \in [0,1] : X(t) \geq u_{\alpha/2}(t) \vee X(t) \leq -u_{\alpha/2}(t)\big) \\
\leq \; & \mathbb{P}\big(\exists\, t \in [0,1] : X(t) \geq u_{\alpha/2}(t)\big) + \mathbb{P}\big(\exists\, t \in [0,1] : X(t) \leq -u_{\alpha/2}(t)\big) \\
\leq \; & \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha.
\end{aligned}$$

[5] see Adler and Taylor (2007), Preface.

This is equivalent to

$$\mathbb{P}\big(\forall\, t \in [0,1] : X(t) \in [-u_{\alpha/2}(t), u_{\alpha/2}(t)]\big) \geq 1-\alpha.$$

Plugging in the definition of $X(t)$, Equation (3.1), yields the desired simultaneous confidence band for our parameter function $\theta(t)$ such that

$$\begin{aligned} &\mathbb{P}\big(\forall\, t \in [0,1] : \theta(t) \in [\hat{\theta}_l(t), \hat{\theta}_u(t)]\big) \geq 1-\alpha, \\ \text{with} \quad &\hat{\theta}_l(t) = \hat{\theta}(t) - u_{\alpha/2}(t)\sqrt{\mathrm{Var}(\hat{\theta}(t))} \\ \text{and} \quad &\hat{\theta}_u(t) = \hat{\theta}(t) + u_{\alpha/2}(t)\sqrt{\mathrm{Var}(\hat{\theta}(t))}. \end{aligned} \tag{3.6}$$

Remark 3.2. Note that the main challenges for the construction of these confidence bands involve the following: We need to find an estimator $\hat{\theta}(t)$ for our parameter function and then determine its variance $\mathrm{Var}(\hat{\theta}(t))$. Especially, solving the Kac-Rice formula in Equation (3.4) requires knowledge about the distribution of $X(t)$. This specifically entails knowledge of, or the ability to consistently estimate, the roughness parameter $\tau(t)$ of the process $X(t)$. We address the complexities of estimating this parameter in detail in Section 4.

Remark 3.3 (**One-sided Confidence Bands**)**.** Certainly, one-sided $(1-\alpha)$-confidence bands can also be constructed by slightly adapting the presented approach. Specifically, by selecting a critical value function $u_\alpha \in U_\alpha([0,1])$ in place of (3.5). This yields Equation (3.6) for the bands $[\hat{\theta}_l(t), \infty)$ or $(-\infty, \hat{\theta}_u(t)]$.

Remark 3.4 (**Hypothesis Tests**)**.** Simultaneous confidence bands are ideally suited for hypothesis testing. For that purpose, consider the hypothesis $\mathcal{H}_0 : \theta(t) = \theta_0(t)$ for all $t \in [0,1]$ which states that a function $\theta(t)$ equals a given function $\theta_0(t)$ uniformly over $[0,1]$. This can be tested on a $(1-\alpha)$-confidence level against the two-sided alternative $\mathcal{H}_1 : \theta(t) \neq \theta_0(t)$ for some $t \in [0,1]$ by rejecting the hypothesis if the function θ_0 is not uniformly contained in the band $[\hat{\theta}_l(t), \hat{\theta}_u(t)]$. We can also test against a one-sided alternative such as $\mathcal{H}_1 : \theta(t) < \theta_0(t)$ for some $t \in [0,1]$ or $\mathcal{H}_1 : \theta(t) > \theta_0(t)$ for some $t \in [0,1]$.

Remark 3.5. The construction could have also been based on the more general version of the Kac-Rice formula (2.7) incorporating a general correction location t_0. This is what Liebl and Reimherr (2023) did in their work. However, as demonstrated in Appendix B, the correction location t_0 has, in fact, no influence on the outcome of the band, at least in the case of linear critical value functions. There is no obvious evidence that it has in other cases, that is, that $\mathbb{E}[\varphi_{u,X}(t_0)]$ is different from $\mathbb{E}[\varphi_{u,X}(0)]$ even though $\varphi_{u,X}(0)$ and $\varphi_{u,X}(t_0)$, $t_0 \in [0,1]$, are both Euler characteristics of the same excursion set (2.2). That is why we will only focus here on the more basic formula (2.4).

3.2. Asymptotic Normal Estimators

Up to this point, we have focused on stochastic processes that are centered and elliptically distributed. However, by also considering processes that are asymptotically elliptical, we broaden the field of potential applications of our construction outlined in Subsection 3.1.
In this section, we discuss asymptotic normality, introducing the relevant notion of convergence along with necessary definitions. Subsequently, we present a concise overview of the parameter functions that possess asymptotically normal estimators, thereby demonstrating the applicability of our procedure.

From now on, we denote a sequence of processes $X_n(t)$ indexed by $n \in \mathbb{N}$, analogous to (3.1),

$$X_n(t) = \frac{\hat{\theta}_n(t) - \theta(t)}{(\mathrm{Var}(\hat{\theta}_n(t)))^{1/2}} \tag{3.7}$$

where the $\hat{\theta}_n(t)$ is an estimator for the parameter function $\theta(t)$ that arises empirically from an i.i.d. sample $\{S_i\}_{i=1}^n$. Of course, the underlying distribution of the sample must depend on the parameter $\theta(t)$.
The following definition of process convergence is adapted to our specific setting where we have C^1-random processes and thus, weak convergence for measurable functions. For the general formulation, refer to Pollard (1990), Definition 9.1, or to Van Der Vaart and Wellner (1996), Definition 1.3.3.

Definition 3.6 (**Process Convergence**)**.** Let $X, (X_n)_n : [0,1] \to \mathbb{R}$, $n \in \mathbb{N}$, be random processes. The sequence $(X_n)_n$ weakly converges to the process X if

$$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$$

for all bounded and uniformly continuous functions $f : C^1[0,1] \to \mathbb{R}$. We denote this by $X_n \xrightarrow{d} X$.

Assumption 3.7. Assume that we have a sequence of estimators $(X_n)_n$ such that, in the sense of Definition 3.6,

$$X_n \xrightarrow{d} X \quad \text{as} \quad n \to \infty,$$

where X satisfies Assumption 2.1 that is $X \in C^1[0,1]$ is elliptically distributed. Thus, X_n of Equation (3.7) is asymptotically Gaussian/elliptical in $\mathcal{C}^1[0,1]$. Further assume that the limiting distribution of V in $X \overset{d}{=} VZ$ (as in Lemma 2.3) is known.

Assumption 3.7 is satisfied in particular for estimators $\hat{\theta}_n$ that are asymptotically normal, i.e. that satisfy

$$\sqrt{n}\big(\hat{\theta}_n(t) - \theta(t)\big) \xrightarrow{d} \mathcal{G}(0, C_\theta), \tag{3.8}$$

where $\mathcal{G}(0, C_\theta)$ is a mean-zero Gaussian process with covariance function C_θ. In order to achieve such asymptotic normality, we require that the estimator is unbiased or that its bias is asymptotically negligible and that the estimate is tight so that convergence in distribution occurs in the strong topology.[6]

The assumption of asymptotic normality, as defined in Equation (3.8), is a quite broad and general one. Asymptotic normality is demonstrated for example in mean function estimation with local polynomial estimators, as shown in Degras (2011), Theorem 1, or Berger et al. (2023), Theorem 3. Also, local polynomial estimators are asymptotically normal in covariance estimation, which is shown for the local linear case by Zhang and Wang (2016), Theorem 3.2, or for general orders by Berger and Holzmann (2024). Such covariance estimators could be restricted to univariate functions, for example by considering the diagonal, i.e., pointwise variance, and thereby suiting our univariate setting (3.8). Moreover, eigenfunctions or eigenvalues possess asymptotic normal estimators, as in Kokoszka and Reimherr (2013), as well as estimators for function-on-scalar regression, as in Reimherr and Nicolae (2014).

3.3. Justification for Kac-Rice Approach in Asymptotic Case

Motivated by the previous subsection, we show in this section that our approach of constructing confidence bands still works in the case of asymptotically elliptical processes. Therefore, we aim to justify that we can still apply the Kac-Rice formula of Theorem 2.9 for Equation (3.4) that is solving the expected number of excursions for a confidence level, even if the stochastic process X

[6] discussed in Choi and Reimherr (2018).

is only asymptotically elliptical and its dispersion function $c(t,s)$, $t,s \in [0,1]$, is unknown but can be estimated consistently.
The content of this subsection, especially the assumptions and Theorem 3.9, is also taken from Liebl and Reimherr (2023).
While this part is again more theoretical, we will see an application of the results here in form of an example for the mean function in Subsection 3.4.

It will be demonstrated below that the solution sets for $u(t)$ in Equation (3.4), i.e. $\mathbb{E}[\varphi_{u,X}(0)] = \alpha/2$, exhibit asymptotically equivalent functions for an asymptotically elliptical process X_n and its respective limiting process X. Before proving this convergence, we make additional assumptions regarding the consistency of the estimation of the dispersion function and normalize the processes in order to achieve pointwise unit dispersion.

Theorem 2.9 further requires the dispersion function $c(t,s) = \mathrm{Cov}(Z(t), Z(s))$ of X for the roughness parameter $\tau(t)^2 = \partial_{12}c(t,t) = \mathrm{Var}(Z'(t))$ which both might be unknown. Therefore, we assume that they can be consistently estimated.

Assumption 3.8. We assume to have a sequence of estimators $\hat{c}_n(t,s)$, $n \in \mathbb{N}$, for the dispersion function $c(t,s)$ of process X satisfying

$$\begin{aligned}
\sup_{t\in[0,1]} |\hat{c}_n(t,t) - c(t,t)| &= o_{\mathbb{P}}(1),\\
\sup_{t\in[0,1]} |\partial_1\hat{c}_n(t,t) - \partial_1 c(t,t)| &= o_{\mathbb{P}}(1),\\
\sup_{t\in[0,1]} |\partial_{12}\hat{c}_n(t,t) - \partial_{12} c(t,t)| &= o_{\mathbb{P}}(1),
\end{aligned}$$

and that $c(t,t) > 0$ for all $t \in [0,1]$.

Notice that if Assumption 3.7 is satisfied, meaning that we know the distribution of V and additionally, the covariance structure of the limiting process X is known, then Assumption 3.8 is in particular satisfied since we can deduce the dispersion function $c(t,s)$ using $\mathrm{Cov}(X(t), X(s)) = \mathbb{E}[V^2]c(t,s)$.

Further, Theorem 2.9 expects X to have a pointwise unit dispersion $c(t,t) = 1$, $t \in [0,1]$. To ensure this condition, we use the consistent estimator $\hat{c}_n(t,t)$ and Slutsky's lemma to standardize the processes X_n and X,

$$\tilde{X}_n(t) := \hat{c}_n^{-1/2}(t,t)X_n(t) \xrightarrow{d} c^{-1/2}(t,t)X(t) = \tilde{X}(t) \quad \text{as } n \to \infty. \tag{3.9}$$

Now, the processes $\tilde{X}_n$ and $\tilde{X}$ have the following parameter,

$$\begin{aligned}
\tilde{\tau}_n(t)^2 = \partial_{12}\tilde{c}_n(t,t) \quad \text{where} \quad \tilde{c}_n(t,s) &= \frac{\hat{c}_n(t,s)}{\sqrt{\hat{c}_n(t,t)\hat{c}_n(s,s)}}\\
\text{and} \quad \tilde{\tau}(t)^2 = \partial_{12}\tilde{c}(t,t) \quad \text{where} \quad \tilde{c}(t,s) &= \frac{c(t,s)}{\sqrt{c(t,t)c(s,s)}}.
\end{aligned} \tag{3.10}$$

It is easy to see that $\tilde{c}_n(t,t) = 1$ and $\tilde{c}(t,t) = 1$. Thus, the assumptions of Theorem 2.9 are fulfilled. Accordingly, we can plug $\tilde{\tau}_n$ and $\tilde{\tau}$ into Formula (2.4) in order to solve Equation (3.4). For practical applications, Liebl and Reimherr (2023) state in their Remark 2 that this rather amounts in constructing the band from standardized data.

The following theorem was first introduced and proven by Liebl and Reimherr (2023) as Theorem 3.2. It essentially states that the solution sets of (3.4) for X_n and X asymptotically contain the same functions. We just added another small assumption on the function ς to ensure that sets (3.12) and (3.13) are non-empty.

Theorem 3.9. Let Assumptions 2.1 and 2.6 hold for $\tilde{X}$ and let Assumptions 3.7 and 3.8 hold as well. Fix $\gamma \in (0,1)$ and define $\tilde{\tau}_n(t)$ and $\tilde{\tau}(t)$, for $t \in [0,1]$, as in (3.10). Assume a critical value function $u \in \mathcal{U}$ for some set $\mathcal{U} \subseteq C^1_{a.e.}[0,1]$ that is convex, compact and contains the constant functions (up to an appropriate threshold to maintain compactness). Consider $\mathbb{E}[\varphi_{u,\tilde{X}}(0)]$ as in (2.4) with a general positive roughness function $\eta \in C[0,1]$ where $\eta(t) > 0$ for all $t \in [0,1]$ and consider a non-negative, real-valued slackness function $\varsigma : \mathcal{U} \mapsto \mathbb{R}_{\geq 0}$ that is continuous and, for uniformly increasing u, monotonically decreasing and converging to 0. Define the function

$$f(u,\eta) := \mathbb{E}[\varphi_{u,\tilde{X}}(0)] + \varsigma(u) \tag{3.11}$$

and the sets

$$S_n := f^{-1}_{\tilde{\tau}_n}(\gamma) = \{u \in \mathcal{U} : f(u,\tilde{\tau}_n) = \gamma\}, \quad n \in \mathbb{N}, \tag{3.12}$$

$$S := f^{-1}_{\tilde{\tau}}(\gamma) = \{u \in \mathcal{U} : f(u,\tilde{\tau}) = \gamma\}. \tag{3.13}$$

Then, we have the following:

1. The sets $(S_n)_n$ and S are non-empty and closed with probability one.
2. If $(u_n)_n$ is any sequence with $u_n \in S_n$, then $f(u_n,\tilde{\tau}) \xrightarrow{\mathbb{P}} \gamma$ as $n \to \infty$.
3. It is $S_n \to S$ almost surely in Hausdorff distance defined by[7]

$$d(S_n,S) = \max\left(\sup_{u_n \in S_n}\inf_{u \in S}\|u_n - u\|, \sup_{u \in S}\inf_{u_n \in S_n}\|u_n - u\|\right).$$

Remark 3.10 (Comments on Theorem 3.9)**.**

- For simplicity, we can imagine that $\varsigma \equiv 0$. This so-called slackness function will be of interest later in Section 5 when taking some fairness constraints into account.
- We can interpret the sets $(S_n)_n$ and S as the sets of functions $u \in \mathcal{U}$ for which the expected number of up-crossings is equal for a given roughness parameter $\tilde{\tau}_n$ (empirical estimate) and $\tilde{\tau}$ (theoretical truth) respectively.

Hence, Theorem 3.9 is the justification for using the construction of Subsection 3.1 in order to calculate confidence bands for a process that is only asymptotically elliptical or Gaussian. It yields that, asymptotically, we will still obtain the intended confidence level.

We present the proof of Theorem 3.9. This is an extended and more detailed version of the original proof given by Liebl and Reimherr (2023).

Proof of Theorem 3.9. Recall the following two results from analysis and probability theory that we are going to use:

(EQ) If $f(u,\tau)$ is a continuous mapping $C^1_{a.e.}[0,1] \times C[0,1] \to \mathbb{R}$ and $\mathcal{U}$ is compact, then the collection of functions $\{f_u(\tau) := f(u,\tau), u \in \mathcal{U}\}$ is also compact which implies the equicontinuity of the set. This means[8]

$$\forall\, \epsilon > 0 \quad \exists\, \delta > 0 \quad \forall\, f_u : \quad \|\tau - \tau'\|_\infty \leq \delta \Rightarrow |f_u(\tau) - f_u(\tau')| \leq \epsilon.$$

(SUB) It is $\tau_n \to \tau$ in probability if and only if every subsequence $\tau_{n(m)}$ has another subsequence $\tau_{n(m_l)}$ that converges almost surely (see Theorem 2.3.2. in Durrett (2019)).

From Assumption 3.8, we have for the functions in (3.10) that $\tilde{\tau}_n \xrightarrow{\mathbb{P}} \tilde{\tau}$. Due to (SUB), we will work with the subsequence $\tilde{\tau}_{n(m_l)}$ that converges almost surely. After showing that results 1, 2 and 3 of the theorem hold almost surely for the subsequence of the subsequence, we can then apply the equivalence (SUB) in the reverse direction to obtain the results in probability.
For notational convenience, we still write $\tilde{\tau}_n \to \tilde{\tau}$.

[7] The norm can for example be the C-norm $\|\cdot\|_\infty$.
[8] Compare to Theorem II.3.4 Werner (2007).

1. The sets $(S_n)_n$ and S are non-empty because we can always find a constant function $u \in \mathcal{U}$ that is mapped to γ under $f_{\tilde{\tau}}$. To make this more clear, consider the problem $f_{\tilde{\tau}}(u) \stackrel{!}{=} \gamma$ for a constant $u(t) \equiv u$. Using Formula (2.5) leads to

$$f_{\tilde{\tau}}(u) = \mathbb{P}(\tilde{X}(0) \geq u) + \frac{\|\tilde{\tau}\|_{L^1}}{2\pi} M_{\mathcal{V}}\left(-\frac{u^2}{2}\right) + \varsigma(u) \stackrel{!}{=} \gamma.$$

   Here, $f_{\tilde{\tau}}$ is continuous, $\lim_{u\to\infty} f_{\tilde{\tau}}(u) = 0$ and $\lim_{u\to 0} f_{\tilde{\tau}}(u) > 1$. Hence, by the intermediate value theorem, we find a u to solve the equation for $\gamma \in (0,1)$.
   The sets are closed because they are the pre-images of a closed set $\{\gamma\}$ under a continuous mapping $f_{\tilde{\tau}_n}$ or $f_{\tilde{\tau}}$ respectively.
2. Due to the equicontinuity (EQ) of the f_u, we have by definition: For any $\epsilon > 0$ there exists a $\delta > 0$ such that for all $u \in \mathcal{U}$ it is $|f_u(\tau) - f_u(\tau')| \leq \epsilon$ if $\|\tau - \tau'\|_\infty \leq \delta$. By (SUB), we use the subsequence where $\tilde{\tau}_n \to \tilde{\tau}$ a.s. so that we find with probability one an integer N such that $\|\tilde{\tau}_n - \tilde{\tau}\|_\infty \leq \delta$ for all $n \geq N$. Thus, we obtain for $n \geq N$ with probability one that

$$|\gamma - f(u_n, \tilde{\tau})| = |f(u_n, \tilde{\tau}_n) - f(u_n, \tilde{\tau})| \leq \epsilon$$

   for any sequence $u_n \in \mathcal{S}_n$. This yields, by using (SUB) again, the claimed convergence in probability.
3. We split the Hausdorff distance in two terms,

$$d_{1,n} = \sup_{u_n \in S_n} \inf_{u \in S} \|u_n - u\| \quad \text{and} \quad d_{2,n} = \sup_{u \in S} \inf_{u_n \in S_n} \|u_n - u\|.$$

   and show the convergence separately.

   a) We assume $d_{1,n} \nrightarrow 0$ and lead this to a contradiction.
   Then with probability one, there exists an $\epsilon > 0$ and infinitely many n such that $d_{1,n} \geq \epsilon$. We consider this subsequence where $d_{1,n} \geq \epsilon$ for all n. Since the sets $\{S_n\}$ and S are compact by being closed subsets of a compact set $\mathcal{U}$, the form of $d_{1,n}$ implies the existence of a sequence

$$u_n \in \mathcal{S}_n \quad \text{with} \quad \|u_n - u\| \geq \epsilon \quad \text{for all} \quad u_n \in \mathcal{S}.$$

   We consider another subsequence of u_n such that $u_n \to w$ that exists due to compactness of S_n. Property 2 implies

$$f(u_n, \tilde{\tau}) \to \gamma \quad \Longrightarrow \quad f(w, \tilde{\tau}) = \gamma.$$

   Thus, $w \in \mathcal{S}$ but at the same time $\|u - w\| \geq \epsilon$. This is a contradiction such that $d_{1,n} \to 0$ is shown.

   b) Again we assume $d_{2,n} \nrightarrow 0$ and lead this to a contradiction.
   With probability one, there exists an $\epsilon > 0$ and infinitely many n such that $d_{2,n} > 2\epsilon$. We consider a subsequence where $d_{2,n} > 2\epsilon$ for all n. Since the sets are compact, by construction of $d_{2,n}$, there exists a sequence

$$w_n \in \mathcal{S} \quad \text{with} \quad \|u_n - w_n\| > 2\epsilon \quad \text{for all} \quad u_n \in \mathcal{S}_n. \tag{3.14}$$

   We consider a further convergent subsequence such that $w_n \to w$ for an $w \in S$ that exists due to compactness of S. Thus, we find an integer N such that $\|w_n - w\| \leq \epsilon$ for all $n \geq N$. This implies that $w \notin \mathcal{S}_n$ for all $n \geq N$ by Equation (3.14) as well as

$$\|u_n - w\| = \|u_n - w_n + w_n - w\| \geq \|u_n - w_n\| - \|w_n - w\| > \epsilon$$

for all $u_n \in \mathcal{S}_n$, $n \geq N$. Thus, w is isolated from all but finitely many $\mathcal{S}_n$ in the sense that there exists a ball of radius ϵ around w that does not intersect with any S_n, $n \geq N$. Hence, on the one side, it is $B_\epsilon(w) \cap S_n = \emptyset$ for $n \leq N$.

For a small $c \in \mathbb{R}$, we have $w + 1c \in \mathcal{U}$ with 1 being the constant function since constant functions are contained in $\mathcal{U}$ and $\mathcal{U}$ is compact. If w is closer to the boundary than c, we just enlarge $\mathcal{U}$ as much as needed. Then, for any $\tilde{\tau}$, the function

$$f_{\tilde{\tau}}(w + 1c) = \mathbb{P}(\tilde{X}(0) \leq w(0) + c) + \mathbb{E}[N_{w+1c,\tilde{X}}[0,1]] + \varsigma(w + 1c) \tag{3.15}$$

is monotonically decreasing with increasing c. This is because the critical value function $w + 1c$ is increasing uniformly so that $\mathbb{P}(\tilde{X}(0) \geq w + 1c)$ is decreasing, the expected number of up-crossings is decreasing and the slackness function ς is decreasing by assumption.
Let ϕ be the density of $\tilde{X}(0)$, i.e. the density of a centered elliptical distribution, which is monotonically increasing on the negative domain and monotonically decreasing on the positive domain. This yields with the difference quotient and the monotony

$$|c|\, \phi(w(0) + |c|) \leq |\mathbb{P}(\tilde{X}(0) \leq w(0) + c) - \mathbb{P}(\tilde{X}(0) \leq w(0))| \leq |c|\, \phi(w(0) - |c|),$$

which means that we can raise/lower the critical value function w uniformly by c and produce a change in γ at least as large in magnitude as the corresponding quantity above since all parts of $f(u, \tilde{\tau})$ are monotone as shown in (3.15).

Recall that $w \in \mathcal{S}$ meaning that $f(w, \tilde{\tau}) = \gamma$ and thus $f(w, \tilde{\tau}_n) \to \gamma$ by (EQ). Hence, for large enough $n \geq N$, it is

$$|f(w, \tilde{\tau}_n) - \gamma| < \delta := \frac{\epsilon}{2}\, \phi\left(u(0) + \frac{\epsilon}{2}\right).$$

At the same time, we can construct $u_1, u_2 \in \mathcal{U}$ setting $u_i = w + c_i 1$, for some $|c_i| \leq \epsilon/2$ such that $|u_i - w| \leq \epsilon$ as well as

$$f(u_1, \tilde{\tau}_n) > \gamma \quad \text{and} \quad f(u_2, \tilde{\tau}_n) < \gamma.$$

Note that $\mathcal{U}$ is convex and $f(u_1 t + (1-t)u_2, \tilde{\tau}_n)$ is continuous in $t \in [0,1]$. Thus, on the other side, there must exist a $u \in B_\epsilon(w) \cap \mathcal{U}$ such that $f(u, \tilde{\tau}_n) = \gamma$ implying $u \in \mathcal{S}_n$. This is a contradiction since no element of $\mathcal{S}_n$ is within an ϵ environment of u. Thus $d_{2,n} \to 0$.

□

3.4. Band for Mean Function

The most evident and probably most relevant application of the theory so far is the calculation of confidence bands for the mean function. For the remainder of this work, we focus exclusively on this estimation task. Liebl and Reimherr (2023) provide a concrete procedure for this case in their Examples 1 and 2, which we describe in the following. Note that we first make a concrete assumption of the distribution of our sample which we relax afterwards in the second part of this subsection.

Gaussian Processes

We consider an i.i.d. sample of $\mathcal{G}(\mu, C)$-distributed processes $\{S_i\}_{i=1}^n$. In order to construct a band for the mean function $\mu(t) = \mathbb{E}[S_i(t)]$, we use the unbiased estimator

$$\hat{\mu}_n(t) = n^{-1} \sum_{i=1}^{n} S_i(t) \tag{3.16}$$

as well as, for the unknown covariance function $C(t,s) = \mathrm{Cov}(S_i(t), S_i(s))$, the unbiased estimator

$$\hat{C}_n(t,s) = (n-1)^{-1} \sum_{i=1}^{n} \big(S_i(t) - \hat{\mu}_n(t)\big)\big(S_i(s) - \hat{\mu}_n(s)\big). \tag{3.17}$$

Building on Equation (3.1) where $\mathrm{Var}(\hat{\mu}) = C(t,t)/n$ and calculating

$$\begin{aligned}\frac{\hat{\mu}_n(t) - \mu(t)}{(\hat{C}_n(t,t)/n)^{1/2}} &= \frac{\frac{1}{n}\sum_{i=1}^n S_i(t) - \mu(t)}{\sqrt{\hat{C}_n(t,t)/n}} \cdot \sqrt{\frac{C(t,t)}{C(t,t)}} \\ &= \frac{\frac{1}{n}\sum_{i=1}^n S_i(t) - \mu(t)}{\sqrt{C(t,t)/n}} \cdot \sqrt{\frac{C(t,t)}{\hat{C}_n(t,t)}} \quad \sim \mathcal{N}(0,1) \cdot \sqrt{\chi^2_{n-1}/(n-1)}\,,\end{aligned}$$

we obtain by definition of the t-distribution

$$X_n(t) := \frac{\hat{\mu}_n(t) - \mu(t)}{(\hat{C}_n(t,t)/n)^{1/2}} \sim t_\nu \qquad \text{with} \quad \nu = n-1, \quad t \in [0,1], \tag{3.18}$$

or in accordance with Lemma 2.3

$$X_n(t) \overset{d}{=} V_n Z(t) \qquad \text{with} \quad V_n \sim \sqrt{\nu/\chi^2_\nu}, \quad Z(t) \sim \mathcal{N}(0,1), \quad t \in [0,1].$$

Now, following the methodology outlined in Subsection 3.1, we may proceed with solving (3.4), that is $\mathbb{E}[\varphi_{u,X}(0)] = \alpha/2$, in order to obtain a critical value function $u_{\alpha/2}$. For this purpose, we use the respective version of the Kac-Rice formula for t-distributed processes given in Corollary 2.13. As mentioned already, the respective roughness parameter function $\tau(t) = (\partial_{12} c(t,t))^{1/2}$ is needed. We deduce $\tau(t)$ in a rather straight forward, theoretical way here, while more details and practical approaches are deferred to Section 4.
The dispersion function $c(t,s)$, $t,s \in [0,1]$, is given by

$$\begin{aligned}&\mathrm{Cov}(X_n(t), X_n(s)) = \mathbb{E}[V_n^2]\mathrm{Cov}(Z(t), Z(s)) = \frac{\nu}{\nu-2} c(t,s), \\ \text{since} \quad &\mathbb{E}[V_n^2] = \mathbb{E}[\nu/\chi^2_\nu] = \int_0^\infty \frac{\nu}{x} f_{\chi^2_\nu}(x)\,dx = \int_0^\infty \frac{\nu x^{\frac{\nu}{2}-1}\exp(-\frac{x}{2})}{x 2^{\nu/2}\Gamma(\nu/2)}\,dx \\ &\qquad = \frac{\nu}{\nu-2}\int_0^\infty \frac{x^{\frac{\nu-2}{2}-1}\exp(-\frac{x}{2})}{2^{(\nu-2)/2}\Gamma((\nu-2)/2)}\,dx = \frac{\nu}{\nu-2},\end{aligned}$$

and naturally satisfies $c(t,t) = 1$. Further, we have

$$\mathrm{Cov}(X_n(t), X_n(s)) = \frac{C(t,s)}{(C(t,t)C(s,s))^{1/2}}.$$

Plugging this into the dispersion function $c(t,s) = \frac{\nu-2}{\nu}\mathrm{Cov}(X_n(t), X_n(s))$ and using (3.17) yields the consistent estimate

$$\hat{c}_n(t,s) = \frac{\nu-2}{\nu} \frac{\hat{C}_n(t,s)}{(\hat{C}_n(t,t)\hat{C}_n(s,s))^{1/2}}.$$

Hence, the roughness parameter function $\tau(t)$ can be estimated consistently by

$$\hat{\tau}(t) = \big(\partial_{12}\hat{c}_n(t,s)\big)^{1/2}. \tag{3.19}$$

Finally, this allows us to solve (3.4), obtain the critical value function $u_{\alpha/2}$ and construct the band

$$[\hat{\mu}_l(t), \hat{\mu}_u(t)] = \hat{\mu}(t) \pm u_{\alpha/2}\sqrt{C(t,t)/n}$$

with the desired $(1-\alpha)$-coverage as described in Subsection 3.1.

Asymptotically Gaussian Processes

Certainly, we can not always assume the sample $\{S_i\}_{i=1}^n$ to be Gaussian. However, when only assuming the sample to be i.i.d., we can employ the methodology for asymptotically Gaussian processes detailed in Subsection 3.2 and 3.3.

We stay with the mean estimator $\hat{\mu}_n(t)$ from (3.16) and covariance estimator $\hat{C}_n$ from (3.17) while not knowing anything about their distribution in the case at hand. However, as in (3.8), we have

$$\sqrt{n}\big(\hat{\mu}_n(t) - \mu(t)\big) \xrightarrow{d} \mathcal{G}(0, C)$$

due to the asymptotic normality of the mean estimator which was shown by Degras (2011), Theorem 1. Equivalently to (3.9), Slutsky's lemma yields that pointwise

$$\tilde{X}_n(t) := \frac{\hat{\mu}_n(t) - \mu(t)}{(\hat{C}_n(t,t)/n)^{1/2}} \xrightarrow{d} \mathcal{N}(0,1), \quad t \in [0,1].$$

This process $\tilde{X}_n$ is exactly the same as the one in (3.18), except that we do consider its asymptotic behaviour. This is consistent, as the t-distribution converges to the standard normal for increasing degrees of freedom. Thus, the limiting process $\tilde{X}$ is standard Gaussian, i.e. has pointwise unit variance, such that its covariance function equals the dispersion function $\mathrm{Cov}(\tilde{X}(t), \tilde{X}(s)) = c(t,s)$. Note that, equivalently to (3.10),

$$\mathrm{Cov}\big(\tilde{X}(t), \tilde{X}(s)\big) = \tilde{C}_n(t,s) = \frac{\hat{C}_n(t,s)}{\sqrt{\hat{C}_n(t,t)\hat{C}_n(s,s)}},$$

which can be used to estimate the roughness parameter function $\hat{\tau}(t) = \big(\partial_{12}\tilde{C}_n(t,s)\big)^{1/2}$. In particular, Assumption 3.7 is satisfied. Moreover, Assumption 3.8 is fulfilled since $\hat{C}_n$ is a consistent estimator for C. Thus, with the background of Theorem 3.9, we can solve (3.4), that is $\mathbb{E}[\varphi_{u,\tilde{X}}(0)] = \alpha/2$, with the Gaussian formula, i.e. Corollary 2.12, in order to obtain a critical value function $u_{\alpha/2}$ with which we can construct the confidence band

$$\big[\hat{\mu}_l(t), \hat{\mu}_u(t)\big] = \hat{\mu}(t) \pm u_{\alpha/2}\sqrt{\hat{C}_n(t,t)/n}$$

for the mean function $\mu(t)$ that satisfies the desired $(1-\alpha)$-coverage asymptotically.

4. Roughness Parameter Function

The expectation $\mathbb{E}[\varphi_{u,X}(0)]$ in Theorem 2.9 is taken over the process X, which typically has an unknown distribution in practical settings. Particularly, the dispersion function $c(t,s)$ of process X is not commonly known or directly observable. However, this function is crucial for calculating $\mathbb{E}[\varphi_{u,X}(0)]$ because its second partial derivative, the roughness parameter function $\tau(t)$, plays a vital role in the Kac-Rice formula (2.4).
Furthermore, while our focus has been on processes with paths in $C^1[0,1]$, in practice, these processes are observed only at discrete time points. Thus, it is valuable to explore methods for estimating the roughness parameter $\tau(t)$, which is the focus of the current section.

The method for estimating $\tau(t)$ significantly depends on how the data is observed, whether with or without errors. For instance, in Liebl and Reimherr (2023), the data is assumed to be noise-free. Conversely, the model considered in Berger et al. (2023) or (2024) incorporates additive errors at every design point.

We will first briefly outline some theory regarding the smoothness of the roughness parameter function and then present both methods, one by Liebl and Reimherr (2023) that is also implemented in the `R`-package `ffscb` and one derived from Berger and Holzmann (2024) using a local polynomial approach. For the latter one, we will present some bandwidth selection procedures and perform some brief analysis. At the end, both approaches are compared.

The code of this and the subsequent sections can be found on GitHub.[9]

4.1. Theoretical Investigations

Following Berger and Holzmann (2024), Chapter 2 and in particular Example 2, and utilizing a result from Azmoodeh et al. (2014), we can deduce the smoothness of a Gaussian process Z from the smoothness of the covariance function. This is interesting because we solely consider continuously differentiable processes in this work and may therefore make assumptions about the smoothness of the covariance function from which also, the roughness parameter is calculated. We will see that the smoother the covariance function is, the greater is the smoothness of the resulting processes.

The following definition of Hölder smoothness classes is adapted to bivariate functions on $[0,1]^2$.

Definition 4.1 (**Hölder Class**). A function $f : [0,1]^2 \to \mathbb{R}$ is Hölder-smooth with order $\gamma > 0$ if for all indices $\beta = (\beta_1, \beta_2)^\top$ with $|\beta| \leq \lfloor \gamma \rfloor =: k$ the derivatives $D^\beta f(t,s) = \partial_1^{\beta_1} \partial_2^{\beta_2} f(t,s)$ exist and if the Hölder-norm given by

$$\|f\|_{\mathcal{H},\gamma} = \max_{|\beta| \leq k} \sup_{(t,s) \in [0,1]^2} \left| D^\beta f(t,s) \right| + \max_{|\beta| = k} \sup_{\substack{(t_1,s_1),(t_2,s_2) \in [0,1]^2 \\ (t_1,s_1) \neq (t_2,s_2)}} \frac{\left| D^\beta f(t_1,s_1) - D^\beta f(t_2,s_2) \right|}{\|(t_1,s_1) - (t_2,s_2)\|_\infty^{\gamma - k}} \quad (4.1)$$

is finite. Define the Hölder class with parameters $\gamma > 0$ and $L > 0$ on $[0,1]^2$ by

$$\mathcal{H}_{[0,1]}(\gamma, L) = \{ f : [0,1]^2 \to \mathbb{R} \mid \|f\|_{\mathcal{H},\gamma} \leq L \}.$$

Azmoodeh et al. (2014), Theorem 1, introduce the following necessary and sufficient condition for Hölder continuity of Gaussian processes.

[9] https://github.com/lisa-drsh/cb_KacRice_fda.git

Theorem 4.2 (**Hölder Continuity of Gaussian Processes**). The Gaussian process X is Hölder continuous of any order $a < H$, i.e.

$$|X(t) - X(s)| \leq C_\epsilon \, |t-s|^{H-\epsilon}, \quad \text{for all } \epsilon > 0,$$

if and only if there exist constants c_ϵ such that

$$d_X(t,s) := \mathbb{E}\big[(X(t) - X(s))^2\big]^{1/2} \leq c_\epsilon \, |t-s|^{H-\epsilon}, \quad \text{for all } \epsilon > 0.$$

We now consider the Gaussian process part Z of an elliptical process X, see Lemma 2.3, and its covariance kernel $c : [0,1]^2 \to \mathbb{R}_{\geq 0}$ for which we employ the following assumption.

Assumption 4.3. Let the covariance function c satisfy $c \in \mathcal{H}_{[0,1]^2}(\gamma, L)$ with $\gamma > 2$.

Hence, in accordance with Definition 4.1, the second partial derivatives of c are known to exist. Further, by selecting the value of $\beta = (1,1)$ in the second summand of (4.1) and utilising the inequality $\|c\|_{\mathcal{H},\gamma} \leq L$, it is possible to derive the upper bound

$$|\partial_{12} c(t,t) - \partial_{12} c(t,s)| \leq L \, |t-s|^{\min(\gamma-2,1)}$$

Note that, since $\gamma > 2$ is assumed, the upper bound declines depending on $|t-s|$. Further, using that Z' is also a centered with covariance kernel $\partial_{12} c$, see Theorem A.2, we calculate

$$\begin{aligned} d^2_{Z'}(t,s) &= \mathbb{E}\big[(Z'(t) - Z'(s))^2\big] = \mathrm{Cov}(Z'(t), Z'(t)) - 2\mathrm{Cov}(Z'(t), Z'(s)) + \mathrm{Cov}(Z'(s), Z'(s)) \\ &= \partial_{12} c(t,t) - 2\partial_{12} c(t,s) + \partial_{12} c(s,s) \\ &\leq |\partial_{12} c(t,t) - \partial_{12} c(t,s)| + |\partial_{12} c(s,s) - \partial_{12} c(t,s)| \\ &\leq 2L \, |t-s|^{\min(\gamma-2,1)}. \end{aligned}$$

Thus, for the process Z', we have that

$$d_{Z'}(t,s) \leq \mathrm{const} \, |t-s|^{\min(\gamma-2,1)/2}. \tag{4.2}$$

Due to Theorem 4.2, we can infer from (4.2) that the process Z' is Hölder continuous of any order less than $\min(\gamma-2,1)/2$ which implies continuity of Z'. Consequently, we have the process $Z \in C^1[0,1]$ and therefore $X \in C^1[0,1]$.

This aligns with the findings presented in Section 1.4 of Azaïs and Wschebor (2009). They demonstrate that if a centered Gaussian process possesses a covariance kernel in C^{2k}, then, given additional conditions on the k-th derivative of the covariance, the paths of the process are almost surely in C^k. In our case, $k = 1$. Thus, the order of differentiability of the sample paths is, roughly speaking, half the order of differentiability of the covariance kernel.
Further research on this topic is done, for example, by Scheuerer (2010), who explored the smoothness of general processes, not limited to Gaussians.

Hence, we see that the smoothness of a process is determined by the smoothness of the covariance kernel, more precisely, continuous differentiability depends on the second derivative of the covariance kernel, i.e., the roughness parameter function. This motivates the definition of the roughness parameter function that we use throughout this work, which is, once again,

$$\tau(t) := \sqrt{\partial_{12} c(t,t)}.$$

4.2. Estimation using R-Package ffscb

In the following, we will describe the method to estimate the roughness parameter function $\tau(t)$ of the elliptical process $X(t)$ that is based on the work of Liebl and Reimherr (2023) and that is also implemented in the respective R-package `ffscb`. Remember that we restricted this work to confidence bands for the mean function, as the upcoming approach is specifically designed to this case.

In the previous section, we already mentioned a quite straight forward procedure to estimate $\tau(t)$, see Equation (3.10) for the asymptotic case or Equation (3.19) for a confidence band for the mean function with a Gaussian sample. There, an available consistent estimate $\hat{c}_n$ for the dispersion function c was plugged into the definition of the roughness parameter function τ in order to obtain the consistent estimate

$$\hat{\tau}(t) = (\partial_{12}\hat{c}_n(t,t))^{1/2}.$$

But, according to Liebl and Reimherr (2023), there is an easier way to estimate $\tau(t)$. This way incorporates the relation $\tau(t)^2 = \mathrm{Var}(Z'(t))$ saying that the squared roughness of an elliptical process X equals the variance of the derived Gaussian process $Z'(t)$ from the pointwise relation $X \overset{d}{=} VZ$. While having $\tau(t) := (\partial_{12}c(t,t))^{1/2}$ defined for the covariance kernel c of the centered Gaussian process Z, the above relation $\partial_{12}c(t,t) = \mathrm{Var}(Z'(t))$ is shown in Theorem A.2.

Practically, this involves standardizing the sample $\{S_i\}_{i=1}^n$ empirically, i.e.

$$\tilde{S}_i(t) = \frac{S_i(t) - \hat{\theta}(t)}{(\hat{C}_{\theta,n}(t,t))}, \quad i = 1, \dots, n,$$

where the estimates $\hat{\theta}$ and $\hat{C}_{\theta,n}$ are the same as in (3.16) and (3.17), and differentiating the paths. This leads to the estimator

$$\hat{\tau}^{\mathrm{ffscb}}(t) = \left(\widehat{\mathrm{Var}}(\tilde{S}_1'(t), \dots, \tilde{S}_n'(t))\right)^{1/2}, \quad t \in [0,1], \tag{4.3}$$

where $\widehat{\mathrm{Var}}$ denotes the pointwise empirical variance.

In the R package `ffscb`, we find the function `tau_fun` that computes the estimate $\hat{\tau}^{\mathrm{ffscb}}$. We outlined the procedure in Algorithm 1.

Algorithm 1: Estimation of the roughness parameter function $\tau(t)$

Input: Observed sample of dimension $p \times n$, where p is the number of design points and n is the sample size.

for *each row/discretization point* **do**
- Standardize the data (mean = 0, sd = 1)

for *each column/observed path* **do**
- Spline interpolation of discretely observed values of each standardized process (`stats::splinefun(method="linear")`),
- Find derivative on given evaluation grid numerically (`pracma::fderiv(method="central")`)

Calculate the standard deviation on all evaluation points

Output: estimated roughness on evaluation grid

4.3. Estimation using Bivariate Local Polynomial Estimator

We present an alternative approach for estimating the roughness parameter function using a local polynomial estimator. For a basic introduction to local polynomial estimators, refer to Chapter 1.6 of Tsybakov (2009). The specialization to the bivariate estimator, which will be presented in this subsection, originates from Berger and Holzmann (2024).

With this approach, we can handle data $(S_{i,j}, t_j)$, $i = 1, \dots, n$, $j = 1, \dots, p$, that arises from n i.i.d. realizations of some underlying process S and is observed with pointwise i.i.d. error. We observe on some fixed grid $\{t_j, j = 1, ..., p\}$ where $t_1 < ... < t_p$ and have n observations at each grid point. As proposed in Berger and Holzmann (2024), we consider the model

$$S_{i,j} := S_i(t_j) = \mu(t_j) + Y_i(t_j) + \varepsilon_{i,j}, \quad i = 1, \dots, n, \quad j \in \{1, \dots, m\}, \tag{4.4}$$

where μ is a deterministic mean function and the $\varepsilon_{i,j}$ are i.i.d. additive errors with mean zero. The processes Y_i are i.i.d. copies of a mean-zero, square-integrable random process Y that determines the covariance structure of our observed data. We denote the underlying covariance kernel with

$$C(t, s) = \mathrm{Cov}(S(t), S(s)) = \mathrm{Cov}(Y(t), Y(s)), \quad t, s \in [0, 1]. \tag{4.5}$$

However, we are primarily interested in the dispersion function c which we will deduce later on from the resulting estimates of the covariance kernel C and its derivatives.
This kernel function C is the objective which we intend to estimate in first instance using the bivariate local polynomial approach, as outlined in the reference Berger and Holzmann (2024). For that purpose, we estimate the empirical covariance kernel on the observation grid with

$$\begin{aligned} z_{j,k;n} := \widehat{\mathrm{Cov}}(S_1(t_j), S_1(t_k)) &= \frac{1}{n-1} \sum_{i=1}^{n} (S_{i,j} - \bar{S}_{n,j})(S_{i,k} - \bar{S}_{n,k}) \\ &= \frac{1}{n-1} \sum_{i=1}^{n} (S_{i,j} S_{i,k} - \bar{S}_{n,j} \bar{S}_{n,k}), \end{aligned}$$

where $\bar{S}_{n,j} = n^{-1} \sum_{i=1}^{n} S_{i,j}$ is the mean at point t_j. This yields the linear estimator

$$\begin{aligned} \hat{C}_n(t, s; h) &:= \sum_{j \neq k}^{p} w_{j,k}(t, s; h) z_{j,k;n} \\ &= \frac{1}{n-1} \sum_{i=1}^{n} \sum_{j \neq k}^{p} w_{j,k}(t, s; h)(S_{i,j} S_{i,k} - \bar{S}_{n,j} \bar{S}_{n,k}), \quad (t, s) \in [0, 1], \end{aligned} \tag{4.6}$$

where $h > 0$ is the bandwidth and $w_{j,k}(t, s; h) := w_{j,k;m}(t, s; h; t_1, \dots, t_p)$ are the weights for the point $(t, s) \in [0, 1]$ that depend on the bandwidth h and the design points $t_1, \dots, t_p$. Unlike Berger and Holzmann (2024), we use all the grid points in $[0, 1]^2$ since our covariance kernel is differentiable but we still omit the diagonal $j = k$ due to bias reasons.

For the local polynomial estimator of order m, we define $P_l : [0, 1]^2 \to \mathbb{R}^{l+1}$ for $l = 1, \dots, m$ by

$$P_l(u_1, u_2) := \left(\frac{u_1^l}{l!}, \frac{u_1^{l-1} u_2}{(l-1)!}, \frac{u_1^{l-2} u_2^2}{(l-2)!2!}, \dots, \frac{u_2^l}{l!} \right),$$

and $U_m : [0, 1]^2 \to \mathbb{R}^{N_m}$ with $N_m := \frac{(m+1)(m+2)}{2}$ by

$$U_m(u_1, u_2) := \big(1, P_1(u_1, u_2), \dots, P_m(u_1, u_2)\big)^\top.$$

Further, we have a non-negative kernel function $K : [0,1]^2 \to [0,\infty)$ and set $K_h(t_1,t_2) := K(t_1/h, t_2/h)$ and $U_h(u_1,u_2) := U_m(u_1/h, u_2/h)$. Then, the optimization problem at a point $(t,s) \in [0,1]^2$ is given by

$$\widehat{\mathrm{LP}}(t,s) := \underset{\vartheta \in \mathbb{R}^{N_m}}{\operatorname{argmin}} \sum_{j \neq k}^{p} \left(z_{j,k;n} - \vartheta^\top U_h \begin{pmatrix} t_j - t \\ t_k - s \end{pmatrix} \right)^2 K_h \begin{pmatrix} t_j - t \\ t_k - s \end{pmatrix}. \tag{4.7}$$

By closer investigating the form of the estimator $\widehat{\mathrm{LP}}(t,s)$, we see that we locally approximate the target $C(t,s)$ with a polynomial of order m, motivated by the Taylor series. In order to demonstrate this, we explicitly write it down for order $m = 2$, the case that is of interest for us since it uses the second order derivatives. We have

$$\begin{aligned} C(t_j,t_k) &\approx C(t,s) + \frac{\partial_1 C(t,s)}{1!}(t_j - t) + \frac{\partial_2 C(t,s)}{1!}(t_k - s) + \frac{\partial_{11} C(t,s)}{2!}(t_j - t)^2 \\ &\quad + 2\left(\frac{\partial_{12} C(t,s)}{2!}(t_j - t)(t_k - s) \right) + \frac{\partial_{22} C(t,s)}{2!}(t_k - s)^2 \\ &= \vartheta^\top U_h \begin{pmatrix} t_j - t \\ t_k - s \end{pmatrix}, \end{aligned} \tag{4.8}$$

where

$$\begin{aligned} \vartheta &= \big(C(t,s),\, \partial_1 C(t,s)h,\, \partial_2 C(t,s)h,\, \partial_{11} C(t,s)h^2,\, \partial_{12} C(t,s)h^2,\, \partial_{22} C(t,s)h^2 \big)^\top, \\ U_h(u_1,u_2) &= \big(1,\, u_1/h,\, u_2/h,\, u_1^2/2h^2,\, u_1 u_2/h^2,\, u_2^2/2h^2 \big)^\top. \end{aligned} \tag{4.9}$$

In Formula (4.7), we consider an averaged version of this over the grid points $\{(t_j,t_k), j,k = 1,...,p \text{ with } j \neq k\}$ with a quadratic distance to minimize and weighted with the kernel K_h. Note that after finding the minimizer ϑ in (4.7), we can directly obtain the estimates of all derivatives of the covariance function C up to the respective order because they correspond to the entries of ϑ as in (4.9).

It remains to present the procedure of finding the solution $\widehat{\mathrm{LP}}(t,s)$ in (4.7) for which we follow Berger and Holzmann (2024), Appendix B. In order to differentiate the Formula (4.7) and set it to zero, we rewrite it to

$$\widehat{\mathrm{LP}}(t,s) = \underset{\vartheta \in \mathbb{R}^{N_m}}{\operatorname{argmin}} \left(-2\vartheta^\top a_{p,h}(t,s) + \vartheta^\top B_{p,h}(t,s)\vartheta \right), \quad (t,s) \in [0,1]^2,$$

with

$$a_{p,h}(t,s) := \frac{1}{(ph)^2} \sum_{j \neq k}^{p} U_h \begin{pmatrix} t_j - t \\ t_k - s \end{pmatrix} K_h \begin{pmatrix} t_j - t \\ t_k - s \end{pmatrix} z_{j,k;n} \quad \in \mathbb{R}^{N_m},$$

and

$$B_{p,h}(t,s) := \frac{1}{(ph)^2} \sum_{j \neq k}^{p} U_h \begin{pmatrix} t_j - t \\ t_k - s \end{pmatrix} U_h \begin{pmatrix} t_j - t \\ t_k - s \end{pmatrix}^\top K_h \begin{pmatrix} t_j - t \\ t_k - s \end{pmatrix} \quad \in \mathbb{R}^{N_m \times N_m}.$$

Hence, we need to solve

$$a_{p,h}(t,s) = B_{p,h}(t,s)\vartheta$$

which leads to an unique solution if $B_{p,h}(t,s)$ is positive definite so that

$$\widehat{\mathrm{LP}}(t,s) = \frac{1}{(ph)^2} \sum_{j \neq k}^{p} B_{p,h}^{-1}(t,s) U_h \begin{pmatrix} t_j - t \\ t_k - s \end{pmatrix} K_h \begin{pmatrix} t_j - t \\ t_k - s \end{pmatrix} z_{j,k;n} \quad \in \mathbb{R}^{N_m}. \tag{4.10}$$

The first entry in (4.10) is the linear estimator $\hat{C}_n(t,s;h)$ for the covariance kernel C at point $(t,s) \in [0,1]^2$. It can be either obtained by

$$\hat{C}_n(t,s;h) = \widehat{\mathrm{LP}}_1(t,s) = \widehat{\mathrm{LP}}(t,s)^\top U_m(0,0)$$

or, equivalently, by Formula (4.6) with the weights

$$w_{j,k}(t,s;h) := \frac{1}{(ph)^2} U^\top \begin{pmatrix} 0 \\ 0 \end{pmatrix} B_{p,h}^{-1}(t,s) U_h \begin{pmatrix} t_j - t \\ t_k - s \end{pmatrix} K_h \begin{pmatrix} t_j - t \\ t_k - s \end{pmatrix}.$$

As mentioned before, in particular, we obtain directly all partial derivatives up to order m from Equation (4.10), namely those listed in the vector ϑ in Equation (4.9). To extract those explicitly, we need to multiply with the respective derivative of U_h. For example, in order to get an estimator for $\partial_{12} C(t,s)$, we calculate $\partial_{12} U_h(u_1,u_2) = (0,0,0,0,1/h^2,0)^\top$ such that

$$\partial_{12}\hat{C}(t,s) = \widehat{\mathrm{LP}}(t,s)^\top \partial_{12} U_h(0,0).$$

Remember that, in order to estimate the roughness parameter $\tau(t) := \sqrt{\partial_{12} c(t,t)}$ which is needed in the Kac-Rice formula (2.4), we seek for the estimation of the dispersion function $c(t,s)$ which is standardized on its diagonal, $c(t,t) = 1$. Thus, analogous to Equations (3.10), we standardize the kernel from (4.5) such that

$$c(t,s) := \frac{C(t,s)}{\sqrt{C(t,t)C(s,s)}}.$$

The second partial derivative $\partial_{12} c(t,s)$ can then be calculated by

$$\begin{aligned}
\partial_{12} c(t,s) &= \partial_1 \left(\partial_2 \frac{C(t,s)}{\sqrt{C(t,t)C(s,s)}} \right) \\
&= \partial_1 \left(\frac{\partial_2 C(t,s)}{\sqrt{C(t,t)C(s,s)}} + \frac{C(t,s)\partial_2 (C(s,s)^{-1/2})}{\sqrt{C(t,t)}} \right) \\
&= \frac{\partial_1 \partial_2 C(t,s)}{\sqrt{C(t,t)C(s,s)}} + \frac{\partial_2 C(t,s)\partial_1 (C(t,t)^{-1/2})}{\sqrt{C(s,s)}} + \frac{\partial_1 C(t,s)\partial_2 (C(s,s)^{-1/2})}{\sqrt{C(t,t)}} \\
&\quad + C(t,s)\partial_2 (C(s,s)^{-1/2})\partial_1 (C(t,t)^{-1/2}) \\
&= \frac{\partial_{12} C(t,s)}{\sqrt{C(t,t)C(s,s)}} + \frac{\partial_2 C(t,s)\partial_1 C(t,t)}{2\sqrt{C(t,t)^3 C(s,s)}} + \frac{\partial_1 C(t,s)\partial_2 C(s,s)}{2\sqrt{C(t,t)C(s,s)^3}} + \frac{C(t,s)\partial_1 C(t,t)\partial_2 C(s,s)}{4\sqrt{C(t,t)^3 C(s,s)^3}}.
\end{aligned}$$

Hence

$$\partial_{12} c(t,t) = \frac{\partial_{12} C(t,t)}{C(t,t)} + \frac{5}{4} \frac{\partial_1 C(t,t)\partial_2 C(t,t)}{C(t,t)^2}. \tag{4.11}$$

As described above, we obtain estimates for all required derivatives here such that we can plug them in and find an estimator

$$\hat{\tau}^{\mathrm{locpol}}(t) = \sqrt{\partial_{12}\hat{c}(t,t)} = \sqrt{\frac{\partial_{12}\hat{C}(t,t)}{\hat{C}(t,t)} + \frac{5}{4} \frac{\partial_1 \hat{C}(t,t)\partial_2 \hat{C}(t,t)}{\hat{C}(t,t)^2}}.$$

Remark 4.4 (**Alternative Approach**)**.** In the case of a Gaussian process, we could also first standardize our data $(S_{i,j}, t_j)$, $i = 1, \ldots, n$, pointwise as done in the R function `tau_fun` by Liebl and Reimherr (2023), see Algorithm 1, in order to obtain the pointwise unit dispersion. Based on this pointwise standardized data, we can apply the local polynomial estimator yielding an estimate $\partial_{12}\hat{C}(t,s)$ for $(t,s) \in [0,1]^2$ that can be plugged into $\hat{\tau}(t) = \partial_{12}\hat{C}(t,t)^{1/2}$ directly, such that we can use the construction as in Subsection 3.1 to build confidence bands for the mean function μ.

4.4. Bandwidth Selection

When using a local polynomial estimator, it is necessary to choose a bandwidth. The optimal bandwidth strongly depends on the underlying process as well as the distribution of the noise. Hence, one can not define one bandwidth that works best in every scenario but need to adaptively evaluate for each use case again.
In this section, we perform a bandwidth analysis for a specific process with known covariance. Using a grid search, we find optimal bandwidths minimizing the supremum distance between estimation and true function.

Inspired by model (5.2) of Degras (2011), we consider the centered Gaussian process

$$Y(t) = \eta_1 \sin(\pi t) + \eta_2 (t - 0.5), \quad t \in [0, 1], \tag{4.12}$$

where $\eta_1, \eta_2 \sim \mathcal{N}(0, 1)$ i.i.d. and from which we sample data according to model (4.4) with zero drift $\mu \equiv 0$ and i.i.d. additive error $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. It has the known covariance kernel

$$C(t, s) = \sin(\pi t) \sin(\pi s) + (t - 0.5)(s - 0.5) \tag{4.13}$$

which is plotted in Figure 1.

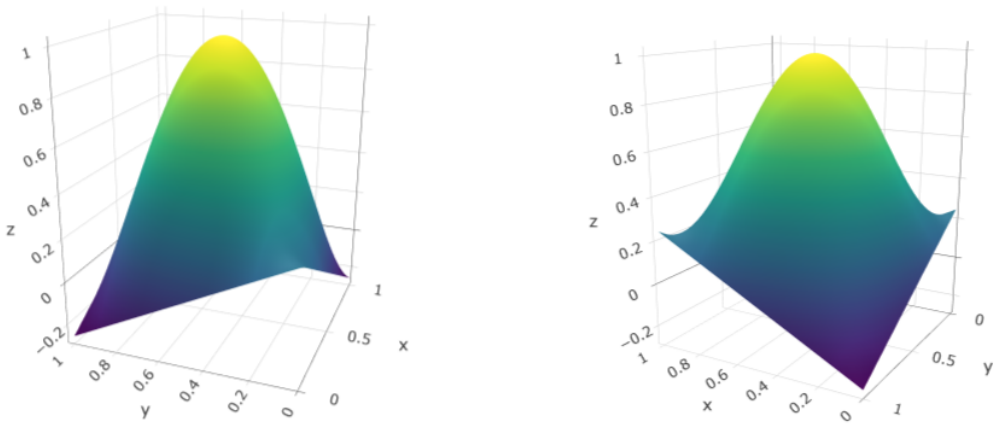

Figure 1: Surface plot of covariance kernel (4.13).

The corresponding derivatives are given by

$$\partial_1 C(t, s) = \pi \cos(\pi t) \sin(\pi s) + (s - 0.5), \tag{4.14}$$

$$\partial_2 C(t, s) = \pi \sin(\pi t) \cos(\pi s) + (t - 0.5), \tag{4.15}$$

$$\partial_{12} C(t, s) = \pi^2 \cos(\pi t) \cos(\pi s) + 1. \tag{4.16}$$

We only consider the diagonal of $[0, 1]^2$ since that is of interest for the estimation of the roughness function $\tau(t)$ as in Equation (4.11). On the diagonal, we have $C(t, t) = \sin(\pi t)^2 + (t - 0.5)^2$, $\partial_1 C(t, t) = \partial_2 C(t, t) = \pi \sin(\pi t) \cos(\pi t) + (t - 0.5)$ and $\partial_{12} C(t, t) = \pi^2 \cos(\pi t)^2 + 1$. Figure 2 shows the form of those functions.

When estimating the roughness function $\tau(t)$ as in (4.11), different derivatives are plugged in. These can be obtained either with the same bandwidth or with individual ones. With a grid search, we aim to find out if the optimal bandwidths for the estimation of the different derivatives coincide.

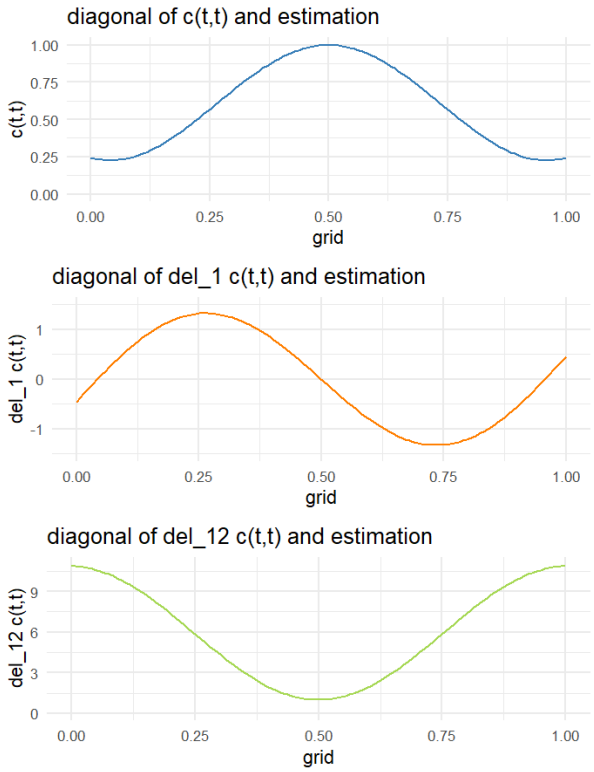

Figure 2: Covariance kernel and derivatives of process (4.12) (note the different scale).

For this purpose, we simulate $n = 100$ samples according to model (4.4) with process $Y(t)$ as in (4.12) and with pointwise errors $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, $\sigma \in \{0, 0.1, 0.3, 0.5\}$, on an equidistant design grid with $p = 120$ points that is also the evaluation grid. For the bandwidths, an equidistant grid on $[0.1, 0.65]$ with step size 0.025 is considered. We average over $N = 1000$ repetitions. Figure 3 shows the resulting supremum error behavior for different standard deviations σ of the pointwise error ε. Similarly, Figure 4 illustrates the error in relation to the size of the original function, i.e. the supremum error was scaled with the maximum value of the true function, in order to asses the relative error that is made. Table 1 summarizes the resulting optimal bandwidths.

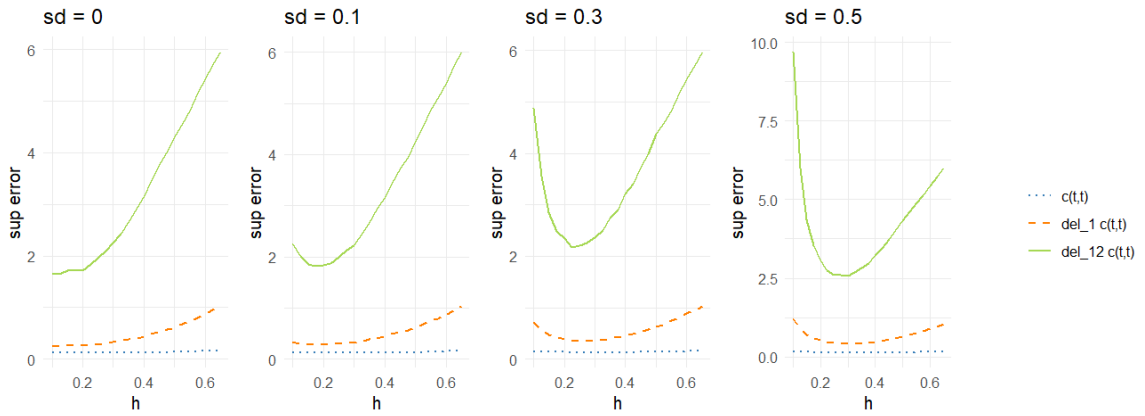

Figure 3: Supremum error for covariance kernel and derivatives for standard deviation of the error.

First of all, Table 1 indicates that we have different optimal bandwidths for the covariance

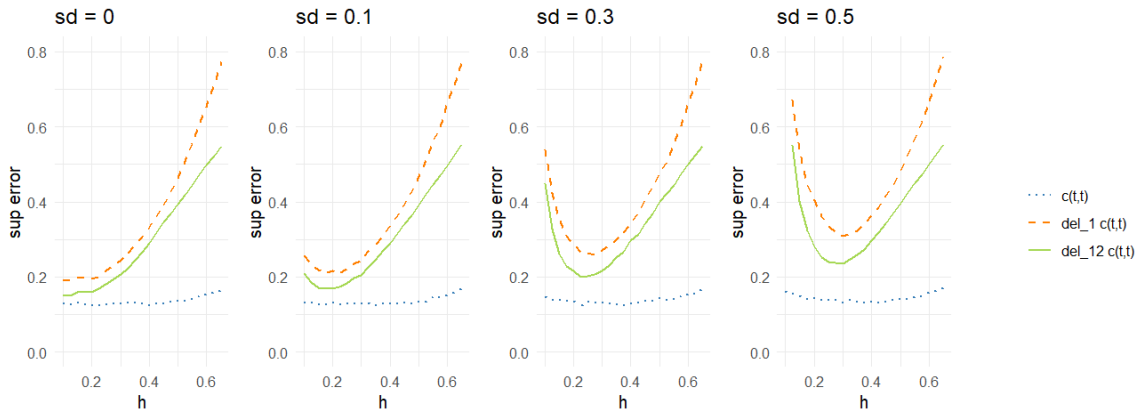

Figure 4: Supremum error (scaled by maximal value of true function) for covariance kernel and derivatives for standard deviation of the error.

σ	$C(t,t)$	$\partial_1 C(t,t)$	$\partial_{12} C(t,t)$
0	0.225	0.125	0.125
0.1	0.350	0.175	0.175
0.3	0.225	0.275	0.225
0.5	0.425	0.300	0.300

Table 1: Optimal bandwidths for different standard deviation of the error, $\sigma \in \{0, 0.1, 0.3, 0.5\}$. Found on equidistant grid on $[0.1, 0.65]$ with step size 0.025.

kernel and its first and second partial derivative. While the covariance kernel itself requires the largest bandwidth to be optimally estimated, the optimal bandwidth declines for higher order of the derivative. This might be explained with a higher variability of the derivatives compared to the kernel function, as can be noted in Figure 2. Also, naturally, the optimal bandwidth increases with stronger noises as more smoothing is required.

Further, in Figure 3, we observe that the estimation error for $\partial_{12} C(t,t)$ is substantially greater than those associated with $\partial_1 C(t,t)$ and $\partial_2 C(t,t)$, with the error belonging to the estimation of $C(t,t)$ being notably the smallest among them. Additionally, the quality of the estimation of $\partial_{12} C(t,t)$ is the most sensitive to the selection of bandwidth, again with the estimation of $C(t,t)$ depending the least on the bandwidth. A similar picture results from the scaled supremum errors in Figure 4. There, the relative error of the first derivative is slightly larger than the error of the second derivative while both are sensitive to the choice of the bandwidth.

It is interesting to investigate the optimal bandwidths again for another design size p and another sample size n since, already shown with simulations by Berger and Holzmann (2024), the optimal bandwidth depends also on those parameters while being more sensitive to p. We repeat the above simulation exact the same way but changing the design size $p = 200$ in one run and changing $n = 200$ in another run. The results can be found in the Appendix C. They draw a very similar picture than Figure 3 and Figure 4.

In conclusion, despite Table 1 advocating for the usage of different bandwidths, and thus separate local polynomial estimations, we suggest to simply use the optimal bandwidth chosen for $\partial_{12} C(t,t)$ due to reasons of computational time. If we would do separate estimations, we would use three times the computational time and by just using the optimal bandwidth for $\partial_{12} C(t,t)$, we only slightly increase the error in supremum norm of $C(t,t)$, $\partial_1 C(t,t)$ and $\partial_2 C(t,t)$ which are relatively small. We decide to use the optimal bandwidth for the estimation of the sec-

ond derivative before the optimal bandwidth for the estimation of the first derivative since the roughness parameter function is computed from the absolute values, see Formula (4.11), such that we must consider Figure 3.

Remark 4.5 (**Bandwidth Selection with Cross-Validation**)**.** In cases where the true roughness function is not known, an alternative approach is required in order to select the optimal bandwidth. This is of particular interest because it would allow the utilization of the *locpol* approach for the estimation of the roughness parameter function in real-world data which generally follows an unknown distribution. Such an alternative approach could, for example, employ a cross-validation technique, as done by Degras (2017) for the mean function. However, for the covariance function, the cross-validation comes with different challenges.

4.5. Empirical Comparison of Estimation Approaches

In this section, we compare the accuracy of the two estimation methods at hand for the roughness parameter function $\tau(t)$ while considering both observations with and without noise. The first method based on the work of Liebl and Reimherr (2023) and described in detail in Subsection 4.2 will be called *ffscb* approach, named after the corresponding `R`-package. The second method following the methodology proposed by Berger and Holzmann (2024) and outlined in Subsection 4.3 will be referred to as the *locpol* approach.

We consider the model (4.4) again with the process (4.12) and $\mu \equiv 0$. The covariance kernel and the partial derivatives are given in (4.13) and (4.14). Plugging these into Formula (4.11) yields the true roughness parameter function $\tau(t)$ in this scenario given by

$$\tau(t)^2 = \frac{\pi^2 \cos(\pi t)^2 + 1}{\sin(\pi t)^2 + (t-0.5)^2} + \frac{5}{4}\left(\frac{\pi \cos(\pi t)\sin(\pi t) + (t-0.5)}{\sin(\pi t)^2 + (t-0.5)^2}\right)^2. \tag{4.17}$$

We simulate $n = 100$ discretely observed paths of the process S with Y as in (4.12) such that

$$S_{i,j} = Y(t_j) + \varepsilon_{ij}, \quad i = 1, ..., n,\ j = 1, ..., p, \tag{4.18}$$

with $p = 120$, the t_j being equidistant on $[0, 1]$ and i.i.d. $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$. Based on this sample, the roughness parameter function $\tau(t)$ is estimated using both the *ffscb* and the *locpol* approach on a fixed evaluation grid with $p_{\text{eval}} = 100$ equidistant grid points. Increasing the noise intensity from $\sigma = 0$ to $\sigma = 0.5$ allows us to compare the behavior of the estimators for such varying scenarios. For the bivariate *locpol* estimator, we used the respective optimal bandwidth chosen in Subsection 4.4.
Figure 5 illustrates the estimated roughness functions across various noise scenarios and the true roughness function, as defined in equation (4.17), using a sample from equation (4.18).
Clearly, in all scenarios shown in Figure 5, the *locpol* approach yields an estimation $\hat{\tau}(t)$ that aligns more closely with the true function τ. Note that the the *ffscb* method, as proposed by Liebl and Reimherr (2023), was not intended to estimate the roughness function τ under noisy conditions. Nonetheless, considering noise in observations is a reasonable assumption for practical applications. Consequently, it is an important finding that the *ffscb* estimation performs poorly in scenarios with noise. Moreover, even in noise-free situations, $\sigma = 0$, the *locpol* approach still outperforms the *ffscb* method in terms of accuracy, as illustrated in the top right of Figure 5.

Note that the shape of the *ffscb* estimations in Figure 5 close to the boundary is not an effect of the estimation being worse at the boundaries. Actually, considering the estimation error proportional to the value of the true roughness function, i.e. $|(\hat{\tau}(t) - \tau(t))/\tau(t)|$, we rather observe a peak around 0.5 instead of high values at the borders, see Figure 6.

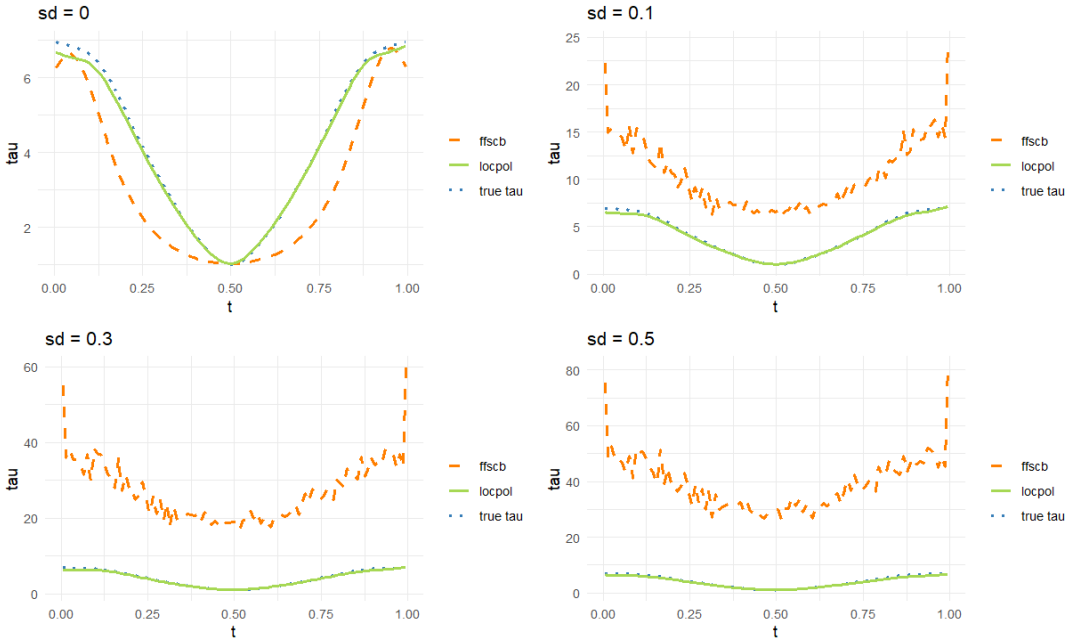

Figure 5: Comparison of roughness estimation for different noise intensities, $\sigma \in \{0, 0.1, 0.3, 0.5\}$, with true roughness parameter function.

We repeat the simulation of the sample (4.18) and estimation of $\tau(t)$ with the two approaches $N = 1000$ times considering different scenarios for the standard deviation of the errors, $\sigma \in \{0, 0.1, 0.3, 0.5\}$, and different design grid sizes, $p \in \{80, 120, 160\}$, while the evaluation grid remains constant with a size of $p_{\text{eval}} = 100$ grid points. Each repetition works just as before estimating the roughness parameter function $\tau(t)$ using the *ffscb* approach as well as the *locpol* approach with the respective optimal bandwidth as chosen in Subsection 4.4.
Measuring the distance from estimation to the true function in supremum norm and averaging over the repetitions yields the following results in Table 2. Table 3 provides the results in relation to the true value of τ, namely we considered the supremum norm of the relative error $\|(\hat{\tau} - \tau)/\tau\|_\infty$.

p	sd = 0		sd = 0.1		sd = 0.3		sd = 0.5	
	locpol	*ffscb*	*locpol*	*ffscb*	*locpol*	*ffscb*	*locpol*	*ffscb*
80	0.792	2.259	0.981	16.543	1.193	52.408	1.135	74.111
	(0.426)	(0.218)	(0.462)	(2.111)	(0.530)	(4.415)	(0.467)	(5.410)
120	0.805	2.267	0.944	16.802	1.043	53.243	1.076	75.867
	(0.438)	(0.222)	(0.449)	(2.165)	(0.448)	(4.731)	(0.425)	(5.217)
160	0.782	2.266	0.875	18.383	1.023	57.548	1.058	80.881
	(0.421)	(0.219)	(0.408)	(2.258)	(0.448)	(5.109)	(0.431)	(5.610)

Table 2: Supremum error of *ffscb* estimation and *locpol* estimation for observations on different grid sizes p and with different errors. In brackets, the standard deviation of the supremum error is displayed.

Once more, it is evident that the *locpol* approach surpasses the *ffscb* method in delivering superior estimations, both in terms of accuracy and precision. This holds true across all examined combinations of design grids and noise levels. While the results for the case sd = 0 are still relatively similar, the errors for the two methods diverge more as the standard deviations increase.

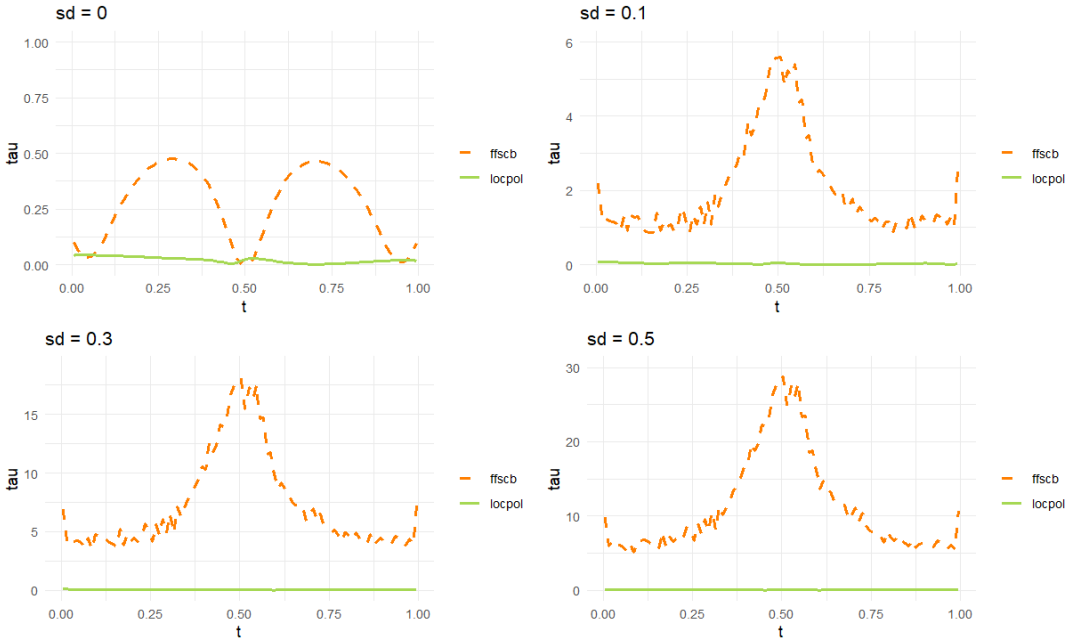

Figure 6: Comparison of pointwise estimation error of roughness estimation proportional to the value of the true roughness function, i.e. $|(\hat{\tau}(t) - \tau(t))/\tau(t)|$, with the same sample from Figure 5.

p	sd = 0		sd = 0.1		sd = 0.3		sd = 0.5	
	locpol	*ffscb*	*locpol*	*ffscb*	*locpol*	*ffscb*	*locpol*	*ffscb*
80	0.129	0.501	0.181	6.578	0.231	20.762	0.202	32.577
	(0.069)	(0.046)	(0.072)	(0.618)	(0.081)	(1.792)	(0.065)	(2.480)
120	0.128	0.502	0.165	6.026	0.202	18.973	0.186	30.063
	(0.069)	(0.045)	(0.068)	(0.605)	(0.073)	(1.592)	(0.059)	(2.199)
160	0.127	0.502	0.154	6.179	0.186	19.534	0.180	30.840
	(0.071)	(0.045)	(0.066)	(0.589)	(0.072)	(1.685)	(0.055)	(2.335)

Table 3: Supremum norm of relative error, i.e. $\|(\hat{\tau} - \tau)/\tau\|_\infty$, of *ffscb* estimation and *locpol* estimation for observations on different grid sizes p and with different errors. In brackets, the standard deviation is displayed.

The *locpol* method shows a small increase in supremum errors, but the *ffscb* method exhibits a very large increase with rising noise intensity. This suggests that the *ffscb* approach is no longer capable of adequately estimating the roughness function in cases involving observations with errors. It would require some more theoretical investigations in order to find out how, exactly, the errors contribute to the estimation of $\hat{\tau}^{\text{ffscb}}$ in (4.3).

Despite knowing that using the supremum norm for measuring the distance between the estimated $\hat{\tau}$ and the true roughness function τ does not particularly advantage the *ffscb* method, as evident in Figure 5 or Figure 6, this metric is nonetheless considered best practice in functional data analysis, e.g. it is also used in Berger and Holzmann (2024). However, for the sake of a fair comparison, it should also be mentioned that the *ffscb* estimation is much faster to compute than the *locpol* approach.

5. Fair Confidence Bands

In this section, we incorporate a fairness constraint into the construction of our confidence bands as proposed by Liebl and Reimherr (2023). In doing so, we aim to obtain bands with a balanced false positive rate.
The false positive rate is generally known from hypothesis testing or machine learning applications[10] and is defined as the ratio of the number of incorrect positive identifications to the total number of actual negative instances. This ratio may be balanced throughout different groups or, as in our case, over a partition of the domain.
In the present context of confidence bands, it corresponds to the probability that a realized path from a process is not covered by the band, even though the band was constructed with its underlying distribution. In order to balance this excursion probability over our domain, the idea is to divide the domain into subintervals and determine a critical value function separately over the subintervals while allowing a proportional share of the confidence level α. Hence, we have an error probability on each subinterval that is proportional to the length of the subinterval. Such confidence bands will be interpretable globally as well as locally.

Rigorous definitions of the above follow in Subsection 5.1. Afterwards, in Subsection 5.2, we present a suitable algorithm to calculate such fair confidence bands in practice. In the last Subsection 5.3, we explore the concept of the price of fairness which quantifies the degree of conservatism resulting from the implementation of a fairness constraint.

As a start, we give a rather simple example that aims to illustrate the advantages of balancing the false positive rate. Although this example does not consider C^1-processes, it effectively demonstrates the underlying principle. The idea for this example arises from Liebl and Reimherr (2023), Chapter 2.1, but calculations were modified.

Motivational Example

Let $S_i \sim \mathcal{G}(\mu, C)$, $i = 1, \dots, n$, be a sample of i.i.d. Gaussian random processes with independent, piecewise constant paths. The paths are constant over two equidistant sections in $[0, 1/3)$, three in $[1/3, 2/3)$ and five in $[2/3, 1]$. Thus, the covariance function $C(t, s) = \mathrm{Cov}(S_1(t), S_1(s))$ is non-stationary and block-diagonal with 10 blocks of different sizes.
More specifically, it is $C(t, s) = c_j$ for $s, t \in [a_{j-1}, a_j)$, $j = 1, 2, 3$, zero otherwise. For $k = 1, 2$, the intervals are equidistant in $[0, 1/3)$, for $k = 3, \dots, 5$, the intervals are equidistant in $[1/3, 2/3)$ and $k = 6, \dots, 10$, the intervals are equidistant in $[2/3, 1]$. We assume the covariance function to be known in this example.

In order to construct a simultaneous confidence band for the unknown mean function $\mu(t) = \mathbb{E}[S_1(t)]$ of the processes, we consider the estimator $\hat{\mu}_n(t) = n^{-1} \sum_{i=1}^{n} S_i(t)$ and the standardized version $X(t) = \sqrt{n}(\hat{\mu}_n(t) - \mu(t))/(C(t, t))^{1/2}$. For every $t \in [0, 1]$, the random variable $X(t)$ is standard normally distributed. Accordingly, we choose a constant upper bound

$$u_{\alpha/2} := q_{1-1/2(1-(1-\alpha)^{1/10})} \tag{5.1}$$

where q denotes the quantile function of the standard normal distribution. Note that this is a different critical value function than in the example of Liebl and Reimherr (2023), Chapter 2.1. We demonstrate below that using (5.1) as a critical value function indeed results in a simultaneous $(1 - \alpha)$-confidence band. Therefore, let t_k, $k = 1, \dots, 10$, be the points in one of the constant parts respectively such that $X(t_k)$, $k = 1, \dots, 10$, are independent and let Z be a

[10] c.f. Hardt et al. (2016).

standard normally distributed random variable. Then using the complement yields

$$\begin{aligned}
&\mathbb{P}\big(\exists\, t \in [0,1] : |X(t)| \geq \big|u_{\alpha/2}\big|\,\big) = 1 - \mathbb{P}\big(\forall\, t \in [0,1] : |X(t)| < \big|u_{\alpha/2}\big|\,\big) \\
&= 1 - \mathbb{P}\Big(\bigcap_{k=1}^{10} \big\{\, |X(t_k)| < \big|u_{\alpha/2}\big| \,\big\}\Big) \overset{\text{iid}}{=} 1 - \mathbb{P}\big(\,|Z| < \big|u_{\alpha/2}\big|\,\big)^{10} \\
&= 1 - \big(1 - (1 - (1-\alpha)^{1/10})\big)^{10} = \alpha.
\end{aligned} \tag{5.2}$$

This is the corrected calculation in comparison to Liebl and Reimherr (2023), Chapter 2.1, which is why we have a different critical value (5.1).
However, the drawback of the constant, non-adaptive upper bound (5.1) becomes apparent when calculating the excursion probability over the subintervals. There, analogous to Equation (5.2), we have the following local confidence levels

$$\begin{aligned}
\mathbb{P}\big(\exists\, t \in [0,1/3) : |X(t)| \geq \big|u_{\alpha/2}\big|\,\big) &= 1 - \mathbb{P}\big(\,|Z| < \big|u_{\alpha/2}\big|\,\big)^{2/10} = 1 - (1-\alpha)^2, \\
\mathbb{P}\big(\exists\, t \in [1/3,2/3) : |X(t)| \geq \big|u_{\alpha/2}\big|\,\big) &= 1 - \mathbb{P}\big(\,|Z| < \big|u_{\alpha/2}\big|\,\big)^{3/10} = 1 - (1-\alpha)^3, \\
\mathbb{P}\big(\exists\, t \in [2/3,1] : |X(t)| \geq \big|u_{\alpha/2}\big|\,\big) &= 1 - \mathbb{P}\big(\,|Z| < \big|u_{\alpha/2}\big|\,\big)^{5/10} = 1 - (1-\alpha)^5.
\end{aligned}$$

Although considering subintervals of the same size, the excursion probabilities, i.e., false positive rates, are unbalanced here. Thus, the band has a varying confidence over the domain.

The solution proposed for this issue by Liebl and Reimherr (2023) is, as mentioned before, to allocate the significance level proportional to the Lebesgue measure of the partition. This will yield a different, in fact adapted, piecewise constant critical value $u_{\alpha/2,j}$ for each partition interval $j = 1, 2, 3$. For this approach, we choose a Šidák correction[11] type of adjustment

$$u_{\alpha/2,j} := q_{1-1/2(1-(1-\alpha/3)^{1/f(j)})}, \quad j = 1,2,3, \tag{5.3}$$

where $f(j)$ is the number of independent piecewise constant paths within a partition interval such that

$$\begin{aligned}
\mathbb{P}\big(\exists\, t \in [0,1/3) : |X(t)| \geq \big|u_{\alpha/2,1}\big|\,\big) &= 1 - \mathbb{P}\big(\,|Z| < \Big|q_{1-1/2(1-(1-\alpha/3)^{1/2})}\Big|\,\big)^2 \\
&= 1 - (1 - (1 - (1-\alpha/3)^{1/2}))^2 = \alpha/3, \\
\mathbb{P}\big(\exists\, t \in [1/3,2/3) : |X(t)| \geq \big|u_{\alpha/2,2}\big|\,\big) &= 1 - \mathbb{P}\big(\,|Z| < \Big|q_{1-1/2(1-(1-\alpha/3)^{1/3})}\Big|\,\big)^3 \\
&= 1 - (1 - (1 - (1-\alpha/3)^{1/3}))^3 = \alpha/3, \\
\mathbb{P}\big(\exists\, t \in [2/3,1] : |X(t)| \geq \big|u_{\alpha/2,3}\big|\,\big) &= 1 - \mathbb{P}\big(\,|Z| < \Big|q_{1-1/2(1-(1-\alpha/3)^{1/5})}\Big|\,\big)^5 \\
&= 1 - (1 - (1 - (1-\alpha/3)^{1/5}))^5 = \alpha/3.
\end{aligned}$$

By connecting the critical value functions $u_{\alpha/2,j}$, $j = 1, 2, 3$, of the subintervals and reversing the standardization, we achieve the same excursion probabilities for the mean function $\mu(t)$. This enables us to construct a simultaneous confidence band $[\hat{\mu}_l(t), \hat{\mu}_u(t)]$ for $t \in [0,1]$ at a global level of $1-\alpha$. Additionally, it allows for local bands $[\hat{\mu}_l(t), \hat{\mu}_u(t)]$ for $t \in [a_{j-1}, a_j]$, where $j = 1, 2, 3$, each maintaining a local level of $1 - \alpha/3$ across three equidistant intervals.

The effect of a fair, adaptive critical value function using $u_{\alpha/2,j}$, $j = 1, 2, 3$, in contrast to the constant, non-adaptive critical value function as in (5.1) is visualized in Figure 7. In both cases, whether constant or fair and adaptive, we have confidence bands for X on an 80% level. It is notable that on subintervals with fewer sections, the bands are narrower; conversely, intervals

[11] see e.g. Lehmann and Romano (2022), Chapter 9.1.2

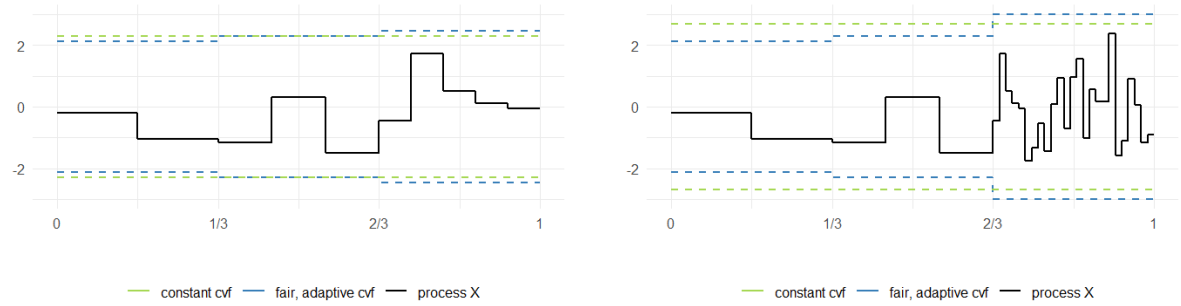

Figure 7: A realized path of X (black) with 80%-confidence bands with critical value functions (5.1) (green) and (5.3) (blue). On the right, the process was adapted to have 25 equidistant sections over $[2/3, 1]$.

with more sections exhibit wider bands. This effect becomes more pronounced with a higher discrepancy in the number of sections between intervals as pointed out by the plot on the right. There, we chose to have 25 of the constant sections instead of 5 in the last third such that the band became wider in order to maintain the level. The reason for this is that the variance of the maximum value of the process increases when the number of sections within a subinterval increases.

5.1. Definition

Returning to processes with C^1-paths, we focus in the remaining section, for the sake of simplicity, on the standardized process X for which the upper confidence bound equals the critical value function u. The fairness as well as excursion probabilities and the confidence level transfer to the parameter process $\theta(t)$ with the respective confidence band, see Proposition 5.4.

Let us specify the formal condition a critical value function $u(t)$ must fulfill in order to be considered as fair. The fundamental idea is that on each subinterval, the significance level should be proportional to the length of the subinterval. Thus, the probability of the process being larger than the boundary, i.e., the critical value, is balanced over the subintervals which we refer to as false positive rate balance. The following definition is taken from Liebl and Reimherr (2023), Definition 2.1 and Lemma 3.2.

Definition 5.1 (**Fair Critical Value Functions**). Let $\{X(t) : t \in [0,1]\}$ be a mean-zero stochastic process with $X \in C^1[0,1]$. Consider a fixed significance level $\alpha \in (0,1)$ and a fixed partition $0 = a_0 < a_1 < \cdots < a_p = 1$, $p \in \mathbb{N}$. The critical value function $u(t)$ is called fair if it allocates the fair proportional shares $\frac{\alpha}{2}(a_j - a_{j-1})$ of the nominal (one-sided) significance level $\alpha/2$ to each local sub-process $\{X(t) : t \in [a_{j-1}, a_j]\}$, $j = 1, 2, \ldots, p$, such that for any $\mathcal{J} \subseteq \{1, 2, \ldots, p\}$,

$$\mathbb{P}\left(\exists\, t \in \bigcup_{j \in \mathcal{J}} [a_{j-1}, a_j] : X(t) \geq u(t)\right) \leq \sum_{j \in \mathcal{J}} \frac{\alpha}{2}(a_j - a_{j-1}). \tag{5.4}$$

The interpretation of a confidence band resulting from such a fair critical value function is possible globally as well as locally. This is because the band will be a global simultaneous confidence band on level α while also providing a $(1-\alpha)(a_j - a_{j-1})$ confidence band on each subinterval $[a_{j-1}, a_j]$, $j = 1, ..., p$, as well as any combination of it.

5.2. Algorithm

We are now taking a step closer to the practical implementation by presenting an algorithm that selects a critical value function $u^*_{\alpha/2} \in \mathcal{U}_{\alpha/2}([0,1])$ satisfying our fairness constraint in Definition 5.1. We denote the function with a star $*$ to indicate the fairness.
Although our approach primarily relies on the algorithm introduced by Liebl and Reimherr (2023), we have implemented some minor adaptations to simplify the algorithm while still obtaining the same results. The reason behind these modifications is discussed subsequently. Additionally, the Appendix D provides extensive justification and empirical studies supporting these changes.

We briefly describe Algorithm 2. To ensure fairness as in Definition 5.1, the domain is divided into p equidistant subintervals, the partitioning being denoted by $0 = a_0 < a_1 < \cdots < a_p = 1$, with the goal of balancing the false positive rates across the subintervals. Going through each interval $[a_j, a_{j+1}]$, the Kac-Rice formula as described in Remark 2.14 is applied.
In the first interval, a critical value function is searched in the one-parameter family of constant functions $\mathcal{U}_{\alpha/2}([0,a_1]) = \{u \in \mathcal{U} : \mathbb{E}[\varphi_{u,X}(0)] = \alpha/2, u \text{ const.}\}$. Afterwards, we search a linear critical value function in the set $\mathcal{U}_{\alpha/2}([a_j,a_{j+1}]) = \{u \in \mathcal{U} : \mathbb{E}[\varphi_{u,X}(a_j)] = \alpha/2, u = u_0 + at\}$ that is still a one-parameter family since the coefficient u_0 is already known from the right interval bound of the previous subinterval $[a_{j-1}, a_j]$ respectively.

This procedure ensures uniqueness of the solution $u^*_{\alpha/2}$ on the one side and continuity on the other side because the linear pieces of the critical value function are compounded continuously. On the grid points $a_0, a_1, \ldots, a_p$ itself, the critical value function will, in general, not be differentiable. This still aligns with all the theory of this work as we assumed that $u \in C^1_{\text{a.e.}}[0,1]$.

Remark 5.2. While Algorithm 2 inherently employs an equidistant fairness partition, it is theoretically capable of functioning identically with any arbitrary partition of the domain. But then, in practice, it is crucial to be cautious with the iterative method. Specifically, if the interval $[a_j, a_{j+1}]$ is chosen too small while the value $u(a_j)$ is comparatively low too, the equation $\mathbb{E}[\varphi_{u,X,[a_j,a_{j+1}]}(a_j)] = \frac{\alpha}{2}(a_{j+1} - a_j)$ may become unsolvable. This occurs because $\mathbb{P}(X(a_j) > u(a_j))$ might already exceed $\frac{\alpha}{2}(a_{j+1} - a_j)$.
Nevertheless, this issue does not arise for an equidistant partition.

Alternative Algorithms

Two additional algorithms can be found in Appendix D, namely Algorithm 3 and 4. The latter one is the one proposed from Liebl and Reimherr (2023). They basically follow the same idea but differ in the way they choose the correction locations, that are the locations where to check whether X is above u. Thus, they use the results of Subsection 2.3.

In more precise terms, Algorithm 4 selects the correction location in an alternating manner. For even indices j, the equation $\mathbb{E}[\varphi_{u,X,[a_{j-1},a_j]}(a_{j-1})] = \frac{\alpha}{2}(a_j - a_{j-1})$ is solved, while for odd indices j, the equation $\mathbb{E}[\varphi_{u,X,[a_{j-1},a_j]}(a_j)] = \frac{\alpha}{2}(a_j - a_{j-1})$ is solved. This naturally complicates the algorithm. However, as demonstrated in Appendix B, the choice of the correction location does not influence the outcome of the Kac-Rice formula for linear critical value functions or rather the root of $\mathbb{E}[\varphi_{u,X}(0)] - \alpha/2 = 0$. Consequently, it can be anticipated that all algorithms will yield the same result, which is indeed the case, as demonstrated in Appendix D.

Therefore, given that all algorithms yield comparable results, we will continue with the straightforward Algorithm 2 presented in this section.

Algorithm 2: Selecting the fair critical value function $u^*_{\alpha/2}$ for correction location always being at left interval endpoint

Input: Provide the roughness parameter function $\tau(t)$, $t \in [0,1]$, set the significance level $\alpha \in (0,1)$ and the number of fairness intervals $p \in \mathbb{N}$.

Initialize the partition $0 = a_0 < \cdots < a_p = 1$ and the critical value function $u_p(t)$ as a piecewise linear function

$$u_p(t) := c_1 + c_2(t - a_1)_+ + \cdots + c_p(t - a_{p-1})_+, \quad t \in [0,1],$$

where $(x)_+ = \max(0, x)$ in order to model $u^*_{\alpha/2}$.

for $j = 1$ **do**

It is $u_p(t) = c_1$ and $u_p'(t) = 0$ for $t \in [0, a_1]$ (constant critical value in first interval). We determine c_1 by solving

$$\mathbb{P}(X(0) \geq u(0)) + \int_0^{a_1} \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\Big(-\frac{1}{2} c_1^2 \Big)\, dt = \frac{\alpha}{2}(a_1 - a_0).$$

for $j = 2, \ldots, p$ **do**

For known $c_1, \ldots, c_{j-1}$, $u_p(t)$ and $u_p'(t)$ for $t \in [a_{j-1}, a_j]$ only depend on c_j the following way

$$u_p(t) = u_p(t, c_j) = c_1 + \sum_{l=1}^{j} c_l (t - a_l)_+ \quad \text{and} \quad u_p'(t) = \sum_{l=1}^{j} c_l.$$

Determine c_j by solving

$$\mathbb{P}(X(a_{j-1}) \geq u(a_{j-1})) + \int_{a_{j-1}}^{a_j} \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\Big(-\frac{1}{2}\Big[u_p(t, c_j)^2 + \frac{u_p'(t, c_j)}{\tau(t)^2} \Big] \Big)\, dt$$
$$- \int_{a_{j-1}}^{a_j} \int_0^{\infty} \frac{u_p'(t, c_j)}{2\pi\tau(t)} M_{\mathcal{V}}'\Big(-\frac{1}{2}\Big[u_p(t, c_j)^2 + \frac{(y + u_p'(t, c_j))^2}{\tau(t)^2} \Big] \Big)\, dy\, dt \quad = \frac{\alpha}{2}(a_j - a_{j-1}).$$

Output: Fair critical value function $u^*_{\alpha/2} := u_p$.

Fairness of the Algorithm

Finally, it remains to show the fairness. As we can observe, the Kac-Rice formula is solved for each subinterval of our partition to adjust the critical value function. In that way, within each subinterval, only the significance level proportional to that subinterval is allowed. Consequently, we obtain a simultaneous confidence band with the specified global confidence level that is fair as well. We state this in the following lemma which was also done before by Liebl and Reimherr (2023) in their Lemma 3.2.

Lemma 5.3 (**Fairness of $u^*_{\alpha/2}$**)**.** Let the conditions of Theorem 2.9 hold and consider a significance level $\alpha \in (0,1)$ and a partition $0 = a_0 < a_1 < \cdots < a_p = 1$, $p \in \mathbb{N}$. Then, the critical value function $u^*_{\alpha/2}$ selected by Algorithm 2 is fair in the sense of Definition 5.1.

Proof. In Algorithm 2, we use the Kac-Rice formula as in Remark 2.14 for every sub-process $\{X(t) : t \in [a_{j-1}, a_j]\}$, $j = 1, 2, \ldots, p$, and solve $\mathbb{E}[\varphi_{u,X,[a_{j-1},a_j]}(a_{j-1})] = \alpha/2(a_j - a_{j-1})$. Thus, for every $j = 1, \ldots, p$, we obtain the inequality

$$\mathbb{P}\big(\exists\, t \in [a_{j-1}, a_j] : X(t) \geq u^*_{\alpha/2}(t)\big) \leq \frac{\alpha}{2}(a_j - a_{j-1})$$

equivalently to the expected Euler characteristic inequality in Theorem 3.1 but on $[a_{j-1}, a_j]$. In order to finally show Equation (5.4) from Definition 5.1 with a general set $\mathcal{J} \subseteq \{1, 2, \dots, p\}$, we only apply Boole's inequality such that

$$\mathbb{P}\Big(\exists\, t \in \bigcup_{j\in\mathcal{J}} [a_{j-1}, a_j] : X(t) \geq u(t)\Big) \leq \sum_{j\in\mathcal{J}} \mathbb{P}\big(\exists\, t \in [a_{j-1}, a_j] : X(t) \geq u^*_{\alpha/2}(t)\big) \leq \sum_{j\in\mathcal{J}} \frac{\alpha}{2}(a_j - a_{j-1}).$$

□

The following proposition corresponds to Proposition 3.2 from Liebl and Reimherr (2023). It states, once again formally, that fair critical value functions result in fair confidence bands.

Proposition 5.4 (**Fair Confidence Bands**)**.** Let the conditions of Theorem 2.9 and Theorem 3.9 hold and let $u^*_{\alpha/2}$ be selected by Algorithm 4 with respect to a given partition and a given α. Let $[\hat{\theta}^*_l, \hat{\theta}^*_u]$ denote the simultaneous confidence band

$$\big[\hat{\theta}^*_l(t), \hat{\theta}^*_u(t)\big] = \Big[\hat{\theta}(t) - u^*_{\alpha/2}(t)\sqrt{\mathrm{Var}(\hat{\theta}(t))},\ \hat{\theta}(t) + u^*_{\alpha/2}(t)\sqrt{\mathrm{Var}(\hat{\theta}(t))}\Big]$$

similar to Equation (3.6). Then it holds that

$$\mathbb{P}\Big(\forall\, t \in \bigcup_{j\in\mathcal{J}} [a_{j-1}, a_j] : \theta(t) \in [\hat{\theta}^*_l(t), \hat{\theta}^*_u(t)]\Big) \geq 1 - \sum_{j\in\mathcal{J}} \alpha(a_j - a_{j-1}) \geq 1 - \alpha$$

for any subset $\mathcal{J} \subseteq \{1, 2, \dots, p\}$.

Proof. This follows directly from Lemma 5.3, i.e.,

$$\mathbb{P}\Big(\exists\, t \in \bigcup_{j\in\mathcal{J}} [a_{j-1}, a_j] : X(t) \geq u(t)\Big) \leq \sum_{j\in\mathcal{J}} \frac{\alpha}{2}(a_j - a_{j-1}) \quad \text{for all} \quad \mathcal{J} \subseteq \{1, 2, \dots, p\}$$

and similar manipulations as in Subsection 3.1 like plugging in $X(t) = (\hat{\theta}(t) - \theta(t))/(\mathrm{Var}(\hat{\theta}(t)))^{1/2}$ and solving for $\theta(t)$. □

5.3. Price of Fairness

The fulfillment of a condition often entails a cost elsewhere. In our case, the cost associated with constructing fair confidence bands manifests as increased width of the bands, resulting in a more conservative estimation, which is generally tried to be avoided in statistics.
We will see that the expensive part contributing to conservatism consists in the correction terms $\mathbb{P}(X(a_j) \geq u(a_j))$, $j = 2, \dots, p$. Thus, the slackness function representing the price of fairness is growing with the size of the partition p.

The following proposition corresponds to Proposition 3.1 from Liebl and Reimherr (2023) but was adapted to fit our Algorithm 2. Further, we slightly changed the claimed inequality to facilitate the interpretability. The proposition discusses the relationship between the width of the simultaneous global bands and the number of fairness intervals

Proposition 5.5 (**Price of Fairness**)**.** The expected Euler characteristic inequality as in Theorem 3.1 when using the fair critical value function $u^*_{\alpha/2}$ determined by Algorithm 2 is

$$\begin{aligned} &\mathbb{P}\big(\exists\, t \in [0,1] : X(t) \geq u^*_{\alpha/2}(t)\big) \leq \frac{\alpha}{2} - \varsigma(u^*_{\alpha/2}), \qquad \text{where} \\ &\varsigma(u^*_{\alpha/2}) = \begin{cases} 0, & \text{if } 1 = p, \\ \sum_{j=0}^{p-1} \mathbb{P}(X(a_j) \geq u^*_{\alpha/2}(a_j)) & \text{if } p \geq 2. \end{cases} \end{aligned} \tag{5.5}$$

We call ς the price of fairness of the critical value function $u^*_{\alpha/2}$.

Proof. Similarly to Remark 2.14, we add a parameter to the Euler characteristic indicating the interval the characteristic is taken on, i.e.,

$$\varphi_{u,X,[a,b]}(a) = \mathbb{P}(X(a) \geq u(a)) + N_{u,X}[a,b].$$

Now, we observe that Algorithm 2 solves on each subinterval

$$\mathbb{E}[\varphi_{u,X,[a_{j-1},a_j]}(a_{j-1})] = \frac{\alpha}{2}(a_j - a_{j-1}), \qquad j = 0,\dots,p-1. \tag{5.6}$$

Note that, in expectation, the number of up-crossings can be summed up by $\mathbb{E}\big[N_{u,X}[a,b]\big] + \mathbb{E}\big[N_{u,X}[b,c]\big] = \mathbb{E}\big[N_{u,X}[a,c]\big]$. Thus, summing (5.6) up for $j = 0,\dots,p-1$ yields $\alpha/2$ for the right side and for the left side

$$\begin{aligned}
&\sum_{j=0}^{p-1} \mathbb{E}\big[\varphi_{u,X,[a_{j-1},a_j]}(a_j)\big] \\
&= \sum_{j=0}^{p-1} \mathbb{P}(X(a_j) \geq u(a_j)) + \mathbb{E}\big[N_{u,X}^{-1}[a_{j-1},a_j]\big] \\
&= \mathbb{P}(X(a_1) \geq u(a_1)) + \mathbb{E}\big[N_{u,X}^{-1}[0,a_1]\big] + \mathbb{E}\big[N_{u,X}[a_1,1]\big] + \sum_{j=0}^{p-1} \mathbb{P}(X(a_j) \geq u(a_j)) \\
&= \mathbb{E}\big[\varphi_{u,X,[0,1]}(a_1)\big] + \varsigma(u),
\end{aligned}$$

where ς is the price of fairness as in Equation (5.5). Hence, with Algorithm 2, we obtain a critical value function $u^*_{\alpha/2}$ that satisfies

$$\mathbb{E}\big[\varphi_{u^*_{\alpha/2},X}(a_1)\big] + \varsigma(u^*_{\alpha/2}) = \frac{\alpha}{2}.$$

Since Theorem 3.1 already yields

$$\mathbb{P}(\exists\, t \in [0,1] : X(t) \geq u^*_{\alpha/2}(t)) \leq \mathbb{E}[\varphi_{u^*_{\alpha/2},X}(a_1)],$$

we actually obtain a critical value function $u(t)$ that satisfies

$$\mathbb{P}(\exists\, t \in [0,1] : X(t) \geq u^*_{\alpha/2}(t)) \leq \frac{\alpha}{2} - \varsigma(u^*_{\alpha/2}).$$

□

The proposition can be interpreted as follows: The upper bound for the excursion probability here is smaller than the required upper bound of $\alpha/2$ for an $(1-\alpha)$-confidence band since $\varsigma(u^*_{\alpha/2}) > 0$. Consequently, the derived critical value function $u^*_{\alpha/2}$ is generally larger than those in $\mathcal{U}_{\alpha/2}([0,1])$ that do not meet a fairness condition. Due to construction (3.6), a larger critical value function $u^*_{\alpha/2}$ results in wider confidence bands. Consequently, adhering to the fairness condition as defined in Definition 5.1 with Algorithm 2 leads to confidence bands that are broader than necessary for a given significance level α.

Remark 5.6. This price of fairness can be compared with the respective theoretical prices of fairness for the other two algorithms, Algorithm 3 and 4, as provided in Appendix D. There is no direct evidence to suggest that any of the algorithms generally results in a lower price of fairness. Consequently, no algorithm can be considered superior to the others in this regard. This also aligns with what could have been expected, given that the resulting bands are identical. However, the prices of fairness do not necessarily coincide since, while having the same number of components, the excursion probabilities are computed at different locations. In practice, the prices of fairness are found to be identical following a brief simulation study, as demonstrated in Appendix D, Table 9 and 10.

6. Simulations: Coverage Probability

In this work, in order to construct confidence bands using the Kac-Rice formula, we have improved the estimation of the parameter function $\tau(t)$ with a new methodology that can also handle observations with errors. Additionally, we introduced a simplified algorithm for calculating fair confidence bands. Now, in this section, it is time to empirically evaluate these developments and compare them, especially to the benchmark which is the methodology proposed by Liebl and Reimherr (2023).

This is achieved by examining the coverage probability, that is, empirically verifying whether the confidence bands maintain the targeted level. More precisely, paths from processes with known distributions are simulated, confidence bands are derived and finally, the proportion of bands that cover the true parameter function is measured.
The parameter function will of course be the mean function here since we only focused on that one and only provided further theory for that one. For simplicity reasons, we set the mean function to zero here, as this does not affect the analysis of the coverage probability.

Regarding the covariance functions, we are going to consider four different scenarios. Three of these are identical to those used in Chapter 4.1 in Liebl and Reimherr (2023) with covariance kernels based on the Matérn covariance

$$C_\theta(t,s) = 0.25^2\big(2^{1-\nu}/\Gamma(\nu)\big)\big(\sqrt{2\nu}\,|t-s|\,\big)^\nu K_\nu\big(\sqrt{2\nu}\,|t-s|\,\big),$$

where Γ is the Gamma function and K_ν is the modified Bessel function of the second kind and where $\nu \leq 0$ controls the roughness of the sample paths. We denote the scenarios

Cov1 Stationary Matérn covariance C_θ with $\nu = 3/2$ (*smooth*),

Cov2 Stationary Matérn covariance C_θ with $\nu = 1/2$ (*rough*),

Cov3 Non-stationary Matérn-type covariance C_θ with $\nu := \nu_{ts} = 2 + \sqrt{\max(t,s)}(1/4 - 2)$ (*smooth to rough*).

Aside form that, we consider the covariance kernel which is well-known from Section 4 already, that is,

Cov4 $C(t,s) = \sin(\pi t)\sin(\pi s) + (t-0.5)(s-0.5)$.

Coverage Probability of Confidence Bands as in Section 3.4

We assess the empirical coverage probability of simultaneous confidence bands for the mean function, constructed according to the methodology outlined in Subsection 3.4.
In order to achieve this, we sample $n = 100$ processes on p_{design} design points and calculate the roughness parameter function $\tau(t)$ with the *ffscb* approach, detailed in Subsection 4.2. By utilizing this $\tau(t)$ and the empirical covariance (3.17), we can compute the confidence band for Gaussian processes as in Corollary 2.12 and for t-processes as in Corollary 2.13. Then, we observe whether the true mean function, namely the zero function, is covered. Note that the critical value function is chosen to be constant here due to the design of the Algorithm 2 (for piecewise linear critical value functions, see the subsequent part). For each scenario, we repeat this procedure $N = 1000$ times, measuring the proportion of bands that do cover the true mean function here over the entire domain.
The results are summarized in Table 4, disaggregated for different confidence level and the distribution for which the Kac-Rice formula was used.

	Gaussian		t-process	
	0.9	0.95	0.9	0.95
Cov1	0.866	0.920	0.897	0.946
Cov2	0.867	0.938	0.904	0.953
Cov3	0.863	0.931	0.895	0.953

(a) $p_{\text{design}} = 120$

	Gaussian		t-process	
	0.9	0.95	0.9	0.95
Cov1	0.870	0.923	0.904	0.952
Cov2	0.861	0.924	0.904	0.959
Cov3	0.864	0.917	0.895	0.948

(b) $p_{\text{design}} = 200$

Table 4: Coverage probability for constant critical value function.

It can be observed that, in general and in accordance with expectations, the confidence bands generated using the Kac-Rice formula for Gaussian processes are tighter than those generated using the formula for t-processes. While the bands for t-processes are close to fulfilling the targeted confidence level, the formula for Gaussian processes led to bands that are rather below the level. Given that the confidence bands are constructed using empirically estimated covariances, it is anyway more appropriate to consider the bands that use the Kac-Rice formula for t-distributed processes.
Furthermore, it can be observed that in both the *smooth* and the *smooth to rough* scenarios, the coverage probability is smaller and almost always below the confidence level. This indicates that the constructed confidence bands are overly restrictive in these cases. The case in which we are able to achieve the correct coverage probability is in *rough* scenarios. However, this can only be achieved by utilising the correct formula for the t-distribution.
The impact of an increase in the observation grid is difficult to quantify. It seems reasonable to posit that there is no impact.
In general, the results of the coverage probability analysis are satisfactory and demonstrate overall good coverage.

Coverage Probability of Fair Confidence Bands according to Section 5

We proceed in a similar way as above, utilising the identical hyperparameter configuration and Algorithm 2, yet with a fairness partition into $p_{\text{fair}} = 6$ and $p_{\text{fair}} = 9$ subintervals. The results are displayed in Table 5.

	Gaussian		t-process	
	0.9	0.95	0.9	0.95
Cov1	0.849	0.915	0.888	0.943
Cov2	0.874	0.930	0.904	0.953
Cov3	0.853	0.923	0.893	0.950

(a) $p_{\text{fair}} = 6$

	Gaussian		t-process	
	0.9	0.95	0.9	0.95
Cov1	0.858	0.933	0.897	0.959
Cov2	0.879	0.940	0.906	0.957
Cov3	0.867	0.930	0.905	0.946

(b) $p_{\text{fair}} = 9$

Table 5: Coverage probability for fair confidence bands.

The observations that can be made here are in agreement with those previously mentioned. Once more, the Kac-Rice formula for t-processes leads to confidence bands with the most suitable coverage probabilities. Furthermore, the *rough* scenario results in wider bands than the other two, while the confidence levels are generally only close to being fulfilled.
In consideration of the size of the fairness partition in comparison, it is observed that for $p_{\text{fair}} = 9$, the coverage probabilities are greater than those for $p_{\text{fair}} = 6$. This is consistent with the theoretical framework of the price of fairness presented in Subsection 5, which suggests that confidence bands become more conservative as the fairness partition increases in size.

Discussion of Coverage Probability using the Different Estimation Approaches of the Roughness Parameter Function Proposed in Section 4

As extensively discussed in Section 4, we have two approaches for estimating the required roughness parameter function, namely the *ffscb* approach, see Subsection 4.2, and the *locpol* approach, introduced in Subsection 4.3. We showed in a simulation study that only the *locpol* approach is able to successfully estimate the roughness parameter function $\tau(t)$ in the case of noisy data while the *ffscb* approach leads to poor estimates in these cases.
Therefore, our aim is to investigate the impact that the choice of roughness parameter estimation has on the confidence bands by examining the covariance probability.

We conduct an equivalent analysis as above, simulating $n = 100$ processes on $p_{\text{design}} = 120$ design points. This time, we solely use the **Cov4** covariance scenario since we know the optimal bandwidths there already, refer to the analysis in Subsection 4.4. We estimate the roughness parameter function $\tau(t)$ with both approaches, *ffscb* as well as *locpol*, calculate the confidence bands for the mean function respectively and obtain the coverage probability is among $N = 1000$ repetitions. Further, we distinguish between estimating the required mean and covariance with the usual estimates as in (3.16) and (3.17) which we denote by *locpol* (a), and estimating them again with local polynomial estimator which we denote by *locpol* (b). This yields the results displayed in Table 6.

	ffscb		*locpol* (a)		*locpol* (b)	
	0.9	0.95	0.9	0.95	0.9	0.95
sd = 0	0.896	0.949	0.921	0.960	0.927	0.965
sd = 0.1	0.938	0.961	0.869	0.917	0.916	0.949
sd = 0.3	0.932	0.957	0.709	0.814	0.879	0.936
sd = 0.5	0.889	0.937	0.462	0.647	0.798	0.879

Table 6: Coverage probability of bands constructed using the roughness parameter function in the ffscb approach and the locpol approach

Surprisingly, the coverage probabilities of the bands constructed with the *ffscb* are much closer to the desired confidence level, even for noisy data. In fact, the coverage probabilities with the *locpol* approach become dramatically worse as the noise intensity increases.
We still note that *locpol* (b) performs better than *locpol* (a) because we used appropriate local polynomial estimates for the mean and covariance there as well but, however, the coverage probability reveals poor behavior also for this methodology.

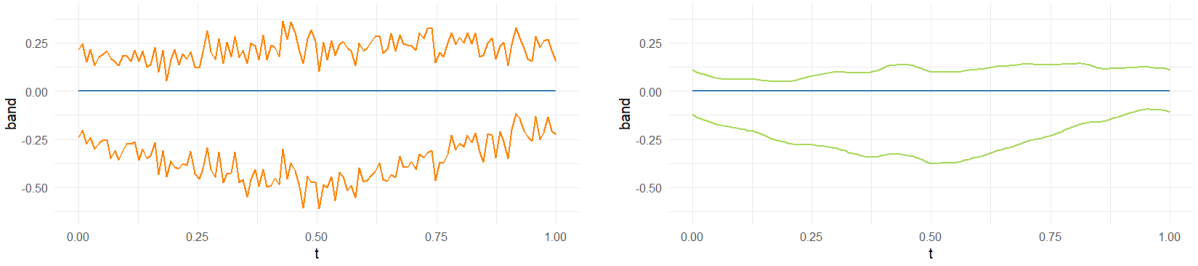

Figure 8: Example of confidence bands around the zero mean function. Left: Confidence band obtained with the *ffscb* approach. Right: Confidence band obtained with the *locpol* (b) approach.

For a better understanding of the problem, Figure 8 shows the shape of the confidence bands

resulting from the two different roughness estimation approaches. For this Figure, one sample with mean zero and covariance **Cov4** is considered with noise intensity $\sigma = 0.5$. It can be observed that the confidence band on the left is more smooth than the one on the right, as would be expected given the application of the smoothing technique. Beside this, it can be speculated that the left confidence band is slightly tighter than the one on the right.

However, the takeaway of this discussion should be that, despite a more accurate estimation of the roughness parameter function $\tau(t)$ with the proposed *locpol* approach in Subsection 4.3, the resulting bands constructed with this more precise estimate do not result in improved coverage probabilities.
We can only assume that in some way the error and resulting uncertainty has not been sufficiently taken into account, as we have smoothed out the errors with our approach.
This definitely is an interesting point and required further research but would go beyond the scope of the current work.

7. Conclusion and Outlook

In this work, we introduced a new local polynomial approach for the estimation of the roughness parameter function that, in contrast to the approach proposed by Liebl and Reimherr (2023), allows to successfully handle data with pointwise additive noise. This was demonstrated empirically in a simulation study, which revealed that the local polynomial approach resulted in greater accuracy and precision in supremum distance, particularly in the presence of noise.

Furthermore, we suggested a simplified algorithm than the one proposed by Liebl and Reimherr (2023) in order to compute fair simultaneous confidence bands. Notably, our algorithm results in the same bands while being more straightforward and using a simpler version of the Kac-Rice formula. The reason for suggesting and implementing a new algorithm is that we corrected the price of fairness which was stated by Liebl and Reimherr (2023) for their algorithm. The rationale for using alternating correction locations was no longer apparent, and thus this idea was omitted, leading to the development of an easier algorithm with a price of fairness of the same magnitude.

Two significant extensions of this work are as follows: Firstly, it would be more applicable to consider processes with paths that are continuous but not necessarily continuously differentiable. This would particularly include the Brownian motion and thus a very wide range of models in application. However, this would necessitate the introduction of a novel definition of up-crossings such that the theory and definitions of Section 2 would be invalid.
Second, the processes could be generalized to a multi-dimensional domain, i.e. considering random fields. The theory of a respective Kac-Rice formula for a constant upper bound u is already available, see Chapter 11.2 in Adler and Taylor (2007).

Prior to this, it would be more profitable in first place to construct bands for different parameter functions than the mean function, using the existing approach. This could involve concretizing our theory for the variance function, eigenfunctions resulting from functional principal component analysis, and so on. This is not an easy problem because either the distribution of $X(t) = (\hat{\theta}(t) - \theta(t))/(\mathrm{Var}(\hat{\theta}(t)))^{1/2}$, $t \in [0,1]$, or a suitable central limit theorem is needed.
In the asymptotic case, the second derivative of the asymptotic covariance kernel is required in order to calculate the roughness parameter function $\tau(t)$. Once asymptotically normal estimators have been obtained, we nearly provided a ready-to-use procedure for estimating $\tau(t)$ using the local polynomial approach from Subsection 4.3, except for no data-driven solution to choose the optimal bandwidth.

Indeed, the bandwidth selection for the local polynomial estimation of the covariance kernel and its derivatives is an interesting direction for further research.
While Berger and Holzmann (2024) have extensively dealt with the optimal bandwidth selection for the covariance function itself, a similar in-depth theoretical analysis of optimal bandwidths for partial derivatives is still lacking. Additionally, a practical approach, for example utilizing cross-validation techniques, requires further development.

Furthermore, it would be interesting to explore critical value functions within a solution set $\mathcal{U}$ that incorporates more than one parameter. In this work, we concentrated on either constant functions $u(t) \equiv u$, solving $\mathbb{E}[\varphi_{u,X}(0)] - \alpha/2 = 0$ for the value u, or on linear functions $u(t) = u_0 + at$ where u_0 was given or implicitly determined, requiring us to find the root of $\mathbb{E}[\varphi_{u,X}(0)] - \alpha/2 = 0$ in terms of a.
Nevertheless, it is conceivable to expand the solution set to encompass functions with more parameters, such as polynomial functions of higher order. In such cases, the roots may no

longer be unique and the number of potential solutions could be infinite. Nevertheless, by applying appropriate constraints, we could effectively select from among the possible solutions. Alternatively, the fair critical value function could be compounded not only continuously but also smoother over the fairness partition. In particular, more flexible functions might offer the advantage of a better fit of the confidence band to the data, provided that sufficient data are available to avoid overfitting.
In these cases, it would require further investigations if the correction location t_0 does also not influence the Kac-Rice formula, i.e., the solution of $\mathbb{E}[\varphi_{u,X}(0)] - \alpha/2 = 0$ for more general critical value functions u than only linear ones.

References

Adler, R. J. and Taylor, J. E. (2007). *Random Fields and Geometry*. Springer Monographs in Mathematics.

Azaïs, J.-M., Bardet, J.-M., and Wschebor, M. (2002). On the tails of the distribution of the maximum of a smooth stationary gaussian process. *ESAIM: Probability and Statistics*, 6:177–184.

Azaïs, J.-M. and Wschebor, M. (2009). *Level Sets and Extrema of Random Processes and Fields*. John Wiley & Sons.

Azmoodeh, E., Sottinen, T., Viitasaari, L., and Yazigi, A. (2014). Necessary and sufficient conditions for hölder continuity of gaussian processes. *Statistics & Probability Letters*, 94:230–235.

Berger, M., Hermann, P., and Holzmann, H. (2023). From dense to sparse design: Optimal rates under the supremum norm for estimating the mean function in functional data analysis. *arXiv:2306.04550*.

Berger, M. and Holzmann, H. (2024). Optimal rates for estimating the covariance kernel in functional data analysis.

Boente, G., Barrera, M. S., and Tyler, D. E. (2014). A characterization of elliptical distributions and some optimality properties of principal components for functional data. *Journal of Multivariate Analysis*, 131:254–264.

Choi, H. and Reimherr, M. (2018). A geometric approach to confidence regions and bands for functional parameters. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 80(1):239–260.

Degras, D. (2017). Simultaneous confidence bands for the mean of functional data. *Wiley Interdisciplinary Reviews: Computational Statistics*, 9(3):e1397.

Degras, D. A. (2011). Simultaneous confidence bands for nonparametric regression with functional data. *Statistica Sinica*, 21(4):1735–1765.

Durrett, R. (2019). *Probability: Theory and Examples*. Cambridge University Press.

Hardt, M., Price, E., and Srebro, N. (2016). Equality of opportunity in supervised learning. *Advances in Neural Information Processing Systems*, 29.

Kokoszka, P. and Reimherr, M. (2013). Asymptotic normality of the principal components of functional time series. *Stochastic Processes and their Applications*, 123(5):1546–1562.

Lalley, S. P. (2011). *Lecture Notes in Brownian Motion and Stochastic Calculus*. University of Chicago - Department of Statistics.

Lehmann, E. L. and Romano, J. P. (2022). *Testing Statistical Hypotheses*. Springer Cham, 4th edition.

Liebl, D. and Reimherr, M. (2023). Fast and fair simultaneous confidence bands for functional parameters. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 85(3):842–868.

Neumann, M. H. and Polzehl, J. (1998). Simultaneous bootstrap confidence bands in nonparametric regression. *Journal of Nonparametric Statistics*, 9(4):307–333.

Pollard, D. (1990). *Empirical Processes: Theory and Applications*. NSF-CBMS Regional Conference Series in Probability and Statistics.

Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis*. Springer, New York, NY, USA, 2nd edition.

Reimherr, M. and Nicolae, D. (2014). A functional data analysis approach for genetic association studies. *The Annals of Applied Statistics*, 8(1):406–429.

Scheuerer, M. (2010). Regularity of the sample paths of a general second order random field. *Stochastic Processes and Their Applications*, 120(10):1879–1897.

Tsybakov, A. B. (2009). *Introduction to Nonparametric Estimation*. Springer Series in Statistics. Springer.

Van Der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer Series in Statistics.

Werner, D. (2007). *Funktionalanalysis*. Springer Spektrum Berlin, Heidelberg.

Zhang, X. and Wang, J.-L. (2016). From sparse to dense functional data and beyond. *The Annals of Statistics*, 44(5):2281–2321.

A. Additional Material for Kac-Rice Formula

A.1. Auxiliary Statements for the Proof

The following lemma is from Liebl and Reimherr (2023), Lemma A.1.

Lemma A.1. Let $f \in C[0,1]$ and suppose there only exists a finite number of zeros $\{t_1, \dots, t_k\}$, that is, $f(t) = 0$ if and only if $t = t_i$ for some $i = 1, \dots, k$. Then, for any $\epsilon > 0$ there exists $h > 0$ such that

$$f^{-1}(-h, h) \subseteq \bigcup_{i=1}^{k} (t_i - \epsilon, t_i + \epsilon). \tag{A.1}$$

The presumption of having finitely many zeros is both reasonable and notably met in our context, where Assumption 2.6 is satisfied. This is further supported by Lemma 6.3.2 in Adler and Taylor (2007).

Proof. We perform the proof by contradiction.
Suppose there exists $\epsilon > 0$ such that Equation (A.1) does not hold for any $h > 0$. In other words, if we let $A_l = f^{-1}(-2^{-l}, 2^{-l})$ for $l = 1, 2, \dots$ and $B_\epsilon = \cup_{i=1}^{k}(t_i - \epsilon, t_i + \epsilon)$, then $A_l \cap B_\epsilon = \emptyset$ or equivalently $A_l \cap B_\epsilon^c \neq \emptyset$.
Selecting an infinite sequence $(x_l)_l$ with $x_l \in A_l \cap B_\epsilon^c$, we have $|f(x_l)| \leq 2^{-l}$ by construction, which implies $f(x_l) \to 0$ as $l \to \infty$.
Since $[0,1]$ is compact and f is continuous by assumption, we have for the limit x of a convergent subsequence $(x_l')_l$ that $f(x) = 0$. But also by construction, it is $|x - t_i| \geq \epsilon$ for all $i = 1, \dots, k$. Hence, we found another zero of the function f that is not contained in $\{t_1, \dots, t_k\}$. This is a contradiction to the assumption that these were the only zeros of the function f. □

Theorem A.2. Let $X = \{X(t), t \in [0,1]\}$ be a mean-zero Gaussian process with $X \in C^1[0,1]$ and covariance function $C(t,s)$, $t, s \in [0,1]$, and let $X' = \{X'(t), t \in [0,1]\}$ be the derivative process. Then, X' is also a mean-zero Gaussian process and its covariance function is given by $\tilde{C}(t,s) = \partial_1 \partial_2 C(t,s)$, $t, s \in [0,1]$.
If X has additionally pointwise constant variance, then $X(t)$ and $X'(t)$ are pointwise uncorrelated and thus independent for all $t \in [0,1]$.

Proof. This proof is inspired by and the first part is mainly conducted analogously to that of Proposition 2.1 in Lalley (2011). First, we define the two-dimensional process

$$DX(t,\epsilon) := \begin{cases} \frac{X(t+\epsilon)-X(t)}{\epsilon} & \text{for } \epsilon \neq 0, \\ X'(t) & \text{for } \epsilon = 0, \end{cases} \qquad t \in [0,1] \text{ and } \epsilon \in \mathbb{R}, \tag{A.2}$$

that is continuous in both parameters because $X(t)$ is continuously differentiable.
To show that $DX(t,\epsilon)$ is Gaussian, we first consider $\epsilon \neq 0$ and note that $X(t+\epsilon)$ and $X(t)$ are Gaussian random variables by assumption. Thus, $DX(t,\epsilon)$ with $\epsilon \neq 0$ is Gaussian as a linear combination of Gaussian random variables. For the case $\epsilon = 0$, we note that by assumption

$$\lim_{\epsilon \to 0} DX(t,\epsilon) = DX(t,0).$$

Since $DX(t,\epsilon)$ was continuous in ϵ, $DX(t,0) = X'(t)$ is Gaussian as the limit of Gaussian random variables. The same way, we show that the derivative process X' is centered. Since X was centered, we have for $\epsilon \neq 0$

$$\mathbb{E}[DX(t,\epsilon)] = \epsilon^{-1}\mathbb{E}[X(t+\epsilon) - X(t)] = 0,$$

and $\mathbb{E}[DX(t,0)] = 0$ due to continuity of $DX(t,\epsilon)$. Hence, the process X' also mean-zero Gaussian under the given conditions on X.

Next, we continue with the covariance structure of X'. We perform a brief calculation for $\epsilon \neq 0$,

$$\begin{aligned}\operatorname{Cov}\big(DX(t,\epsilon), DX(s,\epsilon)\big) &\overset{(A.2)}{=} \frac{1}{\epsilon^2}\operatorname{Cov}\big(X(t+\epsilon) - X(t), X(s+\epsilon) - X(s)\big) \\ &= \frac{1}{\epsilon^2}\big(C(t+\epsilon, s+\epsilon) - C(t+\epsilon, s) - C(t, s+\epsilon) + C(t,s)\big) \\ &= \frac{1}{\epsilon}\bigg(\frac{C(t+\epsilon, s+\epsilon) - C(t+\epsilon, s)}{\epsilon} - \frac{C(t, s+\epsilon) - C(t,s)}{\epsilon}\bigg).\end{aligned}$$

Note that, since process X' is continuous and mean-zero Gaussian, the random variables $\{X'(t)X'(s)\}_{t,s}$ are uniformly integrable and thus, the covariance function $\operatorname{Cov}(X'(t), X'(s)) = \mathbb{E}[X'(t)X'(s)]$ is jointly continuous in $s,t \in [0,1]$. This allows to pull the limit into the expectation such that taking $\epsilon \to 0$ in the above equation leads to

$$\tilde{C}(t,s) := \operatorname{Cov}(X'(t), X'(s)) = \lim_{\epsilon \to 0} \operatorname{Cov}(DX(t,\epsilon), DX(s,\epsilon)) = \partial_1 \partial_2 C(t,s)$$

which is the covariance function of the derivative process X'.

We are moving on to the second part of the claim that is the pointwise independence of X and X'. Considering again $DX(t,\epsilon)$ for $\epsilon \neq 0$, we calculate the covariance of $X(t)$ and $DX(t,\epsilon)$ for a fixed $t \in [0,1]$ which is

$$\begin{aligned}\operatorname{Cov}(X(t), DX(t,\epsilon)) &= \operatorname{Cov}\big(X(t), \frac{X(t+\epsilon) - X(t)}{\epsilon}\big) \\ &= \frac{1}{\epsilon}\big(\operatorname{Cov}(X(t), X(t+\epsilon)) - \operatorname{Cov}(X(t), X(t))\big) \\ &= \frac{1}{\epsilon}\big(C(t, t+\epsilon) - C(t,t)\big)\end{aligned}$$

Now, we apply L'Hôpital's rule and let $\epsilon \to 0$. Since $C(t,s)$ is continuously differentiable such that

$$\lim_{\epsilon \to 0}\left(\frac{\partial}{\partial t}\bigg(C(t, t+\epsilon) - C(t,t)\bigg)\right) = 0,$$

we obtain $\operatorname{Cov}(X(t), X'(t)) = 0$. □

A.2. Special Cases for Arbitrary Correction Location

Equivalently to Subsection 2.2, we want to mention the respective special cases for the generalized Kac-Rice formula (2.7) with arbitrary correction location $t_0 \in [0,1]$ from Theorem 2.17.

Corollary A.3 (Constant Critical Value Function, Arbitrary t_0). We are in the setting of Theorem 2.17 but we let the critical value function be constant, $u(t) \equiv u$. Then, the Kac-Rice formula (2.7) for an elliptical process $X(t)$ and arbitrary correction location $t_0 \in [0,1]$ yields

$$\mathbb{E}[\varphi_{u,X}(t_0)] = \mathbb{P}(X(t_0) \geq u) + \frac{\|\tau\|_{L^1}}{2\pi} M_{\mathcal{V}}\bigg(-\frac{u^2}{2}\bigg).$$

Corollary A.4 (Linear Critical Value Function, Arbitrary t_0). We remain in the setting of Theorem 2.17 but let the critical value function be linear, i.e. $u(t) = u(0) + at$, where a is the

slope. Then, the Kac-Rice formula (2.7) for an elliptical process $X(t)$ and arbitrary correction location $t_0 \in [0,1]$ yields

$$\begin{aligned}\mathbb{E}[\varphi_{u,X}(t_0)] =& \mathbb{P}(X(t_0) \geq u(0) + at_0) + \int_0^1 \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\bigg(-\frac{1}{2}\bigg[(u(0)+at)^2 + \frac{a^2}{\tau(t)^2}\bigg]\bigg)\, dt \\ &+ \int_0^{t_0}\int_0^{\infty} \frac{a}{2\pi\tau(t)} M'_{\mathcal{V}}\bigg(-\frac{1}{2}\bigg[(u(0)+at)^2 + \frac{(y-a)^2}{\tau(t)^2}\bigg]\bigg)\, dy\, dt \\ &- \int_{t_0}^1\int_0^{\infty} \frac{a}{2\pi\tau(t)} M'_{\mathcal{V}}\bigg(-\frac{1}{2}\bigg[(u(0)+at)^2 + \frac{(y+a)^2}{\tau(t)^2}\bigg]\bigg)\, dy\, dt.\end{aligned}$$

Corollary A.5 (**Gaussian Processes, Arbitrary** t_0)**.** In the setting of Theorem 2.17, we assume the process $X(t)$ to be mean-zero Gaussian and let $\mathrm{Var}(X(t)) = \sigma$. Then, the Kac-Rice formula (2.7) with adaptive critical value function $u(t)$ yields

$$\begin{aligned}\mathbb{E}[\varphi_{u,X}(t_0)] =& \Phi\bigg(\frac{-u(t_0)}{\sigma}\bigg) + \int_0^1 \frac{\tau(t)}{2\pi} \exp\bigg(-\frac{1}{2\sigma^2}\bigg[u(t)^2 + \frac{u'(t)^2}{\tau(t)^2}\bigg]\bigg)\, dt \\ &+ \int_0^{t_0} \frac{u'(t)}{\sqrt{2\pi\sigma^2}} \exp\bigg(-\frac{u(t)^2}{2\sigma^2}\bigg)\Phi\bigg(\frac{u'(t)}{\sigma\tau(t)}\bigg)\, dt \\ &- \int_{t_0}^1 \frac{u'(t)}{\sqrt{2\pi\sigma^2}} \exp\bigg(-\frac{u(t)^2}{2\sigma^2}\bigg)\Phi\bigg(\frac{-u'(t)}{\sigma\tau(t)}\bigg)\, dt, \qquad t_0 \in [0,1].\end{aligned}$$

For a linear critical value function, $u(t) = u(0) + at$, we obtain

$$\begin{aligned}\mathbb{E}[\varphi_{u,X}(t_0)] =& \Phi\bigg(\frac{-(u(0)+at_0)}{\sigma}\bigg) + \int_0^1 \frac{\tau(t)}{2\pi} \exp\bigg(-\frac{1}{2\sigma^2}\bigg[(u(0)+at)^2 + \frac{a^2}{\tau(t)^2}\bigg]\bigg)\, dt \\ &+ \int_0^{t_0} \frac{a}{\sqrt{2\pi\sigma^2}} \exp\bigg(-\frac{(u(0)+at)^2}{2\sigma^2}\bigg)\Phi\bigg(\frac{a}{\sigma\tau(t)}\bigg)\, dt \\ &- \int_{t_0}^1 \frac{a}{\sqrt{2\pi\sigma^2}} \exp\bigg(-\frac{(u(0)+at)^2}{2\sigma^2}\bigg)\Phi\bigg(\frac{-a}{\sigma\tau(t)}\bigg)\, dt, \qquad t_0 \in [0,1].\end{aligned} \tag{A.3}$$

For a constant critical value function, $u(t) \equiv u$, we obtain

$$\mathbb{E}[\varphi_{u,X}(t_0)] = \Phi\bigg(\frac{-u}{\sigma}\bigg) + \frac{\|\tau\|_{L^1}}{2\pi} \exp\bigg(-\frac{u^2}{2\sigma^2}\bigg), \qquad t_0 \in [0,1].$$

Corollary A.6 (t**-Processes, Arbitrary** t_0)**.** In the setting of Theorem 2.17, we assume the process $X(t)$ to be a t-process with ν degrees of freedom such that $\mathcal{V} \sim \chi^2_\nu/\nu$. Then, the Kac-Rice formula (2.7) with adaptive critical value function $u(t)$ yields for $t_0 \in [0,1]$

$$\begin{aligned}\mathbb{E}[\varphi_{u,X}(t_0)] =& F_{t_\nu}(-u(t_0)) + \int_0^1 \frac{\tau(t)}{2\pi}\bigg(1 + \frac{u(t)^2}{\nu} + \frac{u'(t)^2}{\tau(t)^2\nu}\bigg)^{-\nu/2} dt \\ &+ \int_0^{t_0} \frac{u'(t)}{2\pi\tau(t)}\bigg(1 + \frac{u(t)^2}{\nu}\bigg)^{-\nu/2-1} \frac{\Gamma((\nu-1)/2)\sqrt{(\nu+1)\pi}a(t)}{\Gamma((\nu+2)/2)} F_{t_{\nu+1}}\bigg(\frac{u'(t)}{a(t)}\bigg)\, dt \\ &- \int_{t_0}^1 \frac{u'(t)}{2\pi\tau(t)}\bigg(1 + \frac{u(t)^2}{\nu}\bigg)^{-\nu/2-1} \frac{\Gamma((\nu-1)/2)\sqrt{(\nu+1)\pi}a(t)}{\Gamma((\nu+2)/2)} F_{t_{\nu+1}}\bigg(\frac{-u'(t)}{a(t)}\bigg)\, dt,\end{aligned}$$

where F_{t_ν} is the cumulative distribution function of the t-distribution with ν degrees of freedom, Γ is the gamma function and $a(t)^2 := \nu\tau(t)^2(1 + u(t)^2/\nu)/(\nu+1)$.
For a constant critical value function, $u(t) \equiv u$, we obtain for $t_0 \in [0,1]$

$$\mathbb{E}[\varphi_{u,X}(t_0)] = F_{t_\nu}(-u) + \frac{\|\tau\|_{L^1}}{2\pi}\bigg(1 + \frac{u^2}{\nu}\bigg)^{-\nu/2}.$$

Remark A.7. In the setting of Theorem 2.17, we consider a subinterval $[a_1, a_2] \subseteq [0,1]$ and the Euler characteristic

$$\varphi_{u,X,[a_1,a_2]}(t_0) = \mathbb{P}(X(t_0) \geq u(t_0) + N_{u,X}^{-1}[a_1, t_0] + N_{u,X}[t_0, a_2], \quad t_0 \in [a_1, a_2],$$

that is only taken on this subinterval. Then, equivalently to the Kac-Rice formula (2.7), we obtain for $t_0 \in [a_1, a_2]$

$$\begin{aligned}\mathbb{E}[\varphi_{u,X,[a_1,a_2]}(t_0)] = &\mathbb{P}(X(t_0) \geq u(t_0)) + \int_{a_1}^{a_2} \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\bigg(- \frac{1}{2}\bigg[u(t)^2 + \frac{u'(t)^2}{\tau(t)^2}\bigg]\bigg)\, dt \\ &+ \int_{a_1}^{t_0}\int_0^{\infty} \frac{u'(t)}{2\pi\tau(t)} M'_{\mathcal{V}}\bigg(- \frac{1}{2}\bigg[u(t)^2 + \frac{(y-u'(t))^2}{\tau(t)^2}\bigg]\bigg)\, dy\, dt \\ &- \int_{t_0}^{a_2}\int_0^{\infty} \frac{u'(t)}{2\pi\tau(t)} M'_{\mathcal{V}}\bigg(- \frac{1}{2}\bigg[u(t)^2 + \frac{(y+u'(t))^2}{\tau(t)^2}\bigg]\bigg)\, dy\, dt.\end{aligned}$$

B. Dependence of Band on Correction Location

During the main part of this work, we used the Kac-Rice formula from Theorem 2.9 which is Theorem 2.17 but with fixed correction location $t_0 = 0$. In the following, we show that, in fact, the confidence band does not depend on the choice of the correction location $t_0 \in [0,1]$ in Theorem 2.17. This means that solving

$$\mathbb{E}[\varphi_{u,X}(t_0)] \stackrel{!}{=} \alpha/2 \tag{B.1}$$

for $u \in \mathcal{U}$ leads to the same solution set $\mathcal{U}_{\alpha/2}([0,1])$ as in (3.5) for all $t_0 \in [0,1]$.

Remark B.1. The only case where we can tell for sure that the choice of t_0 has no influence is if we restrict ourselves to constant critical value functions as in Corollary 2.10. There, the value of $\mathbb{E}[\varphi_{u,X}(t_0)]$ does clearly not depend on t_0 at all such that equating with $\alpha/2$ and solving for $u(t) \equiv u$ will yield a unique solution independent of t_0.

So for now, we let $\mathcal{U}$ be an one-parameter family, containing linear critical value functions $u(t) = u(0) + at$ with fixed $u(0)$. Thus, the slope a is the parameter to determine. This construction suits Algorithms 2, 3 and 4. Thus, finding the solution set for Equation (B.1), that is finding the zero of

$$f(a) := \mathbb{E}[\varphi_{u,X}(t_0)] - \alpha/2 \stackrel{!}{=} 0, \tag{B.2}$$

requires numerical investigations due to the complexity of the form of the Kac-Rice formula (2.7). We use a roughness parameter function τ that results directly from a given covariance kernel in order to be as precise as possible and avoid estimation errors from that. Further, we use a Gaussian process X and the respective formula in Equation (A.3).

Remark B.2. It is not possible to find a solution for a for every choice of $u(0)$ in the above experiment. This is because all the summands $\mathbb{E}[\mathbb{1}_{X(t_0)\geq u(t_0)}]$, $\mathbb{E}[N_{u,X}^{-1}[0,t_0]]$, $\mathbb{E}[N_{u,X}[t_0,1]]$ are positive and one of them might already be larger than $\alpha/2$.

Indeed, the roots for a in Equation (B.2) are exactly the same for any choice of $t_0 \in [0,1]$ when numerically integrating with sufficient precision. The entire function f is even exactly the same. This applies to several choices of $u(0)$ and for the covariance scenarios **Cov1**, **Cov2**, **Cov3**, **Cov4** detailed in Section 6.

C. Additional Results for Bandwidth Selection in Roughness Parameter Estimation

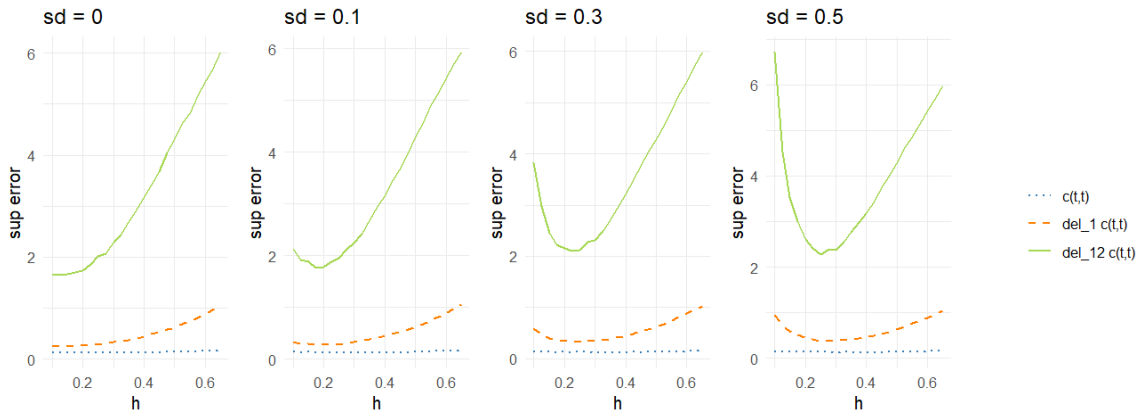

Figure 9: Supremum error for covariance kernel and derivatives for standard deviation of the error. Parameter setting $n = 100$, $p = 200$, $p_{\text{eval}} = 120$.

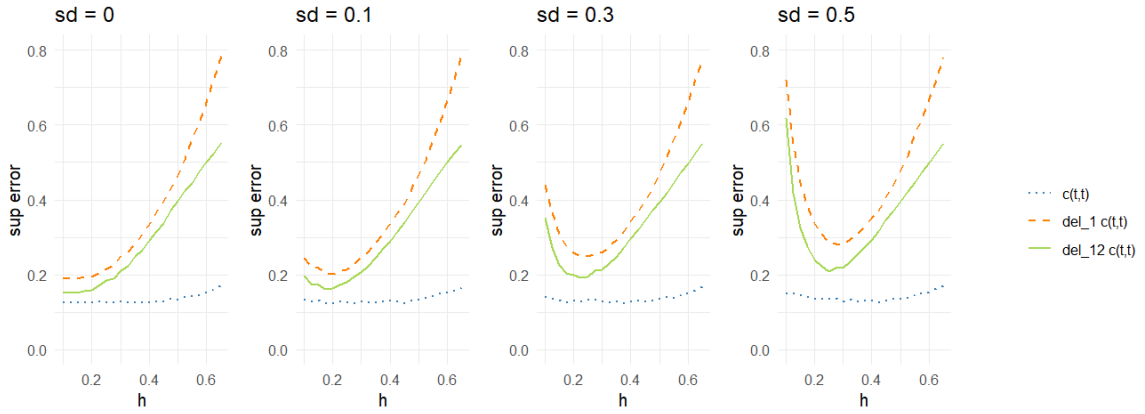

Figure 10: Supremum error (scaled by maximal value of true function) for covariance kernel and derivatives for standard deviation of the error. Parameter setting $n = 100$, $p = 200$, $p_{\text{eval}} = 120$.

σ	$C(t,t)$	$\partial_1 C(t,t)$	$\partial_{12} C(t,t)$
0	0.325	0.150	0.125
0.1	0.200	0.200	0.175
0.3	0.375	0.225	0.225
0.5	0.425	0.300	0.250

Table 7: Optimal bandwidths for different standard deviations of the error, $\sigma \in \{0, 0.1, 0.3, 0.5\}$. Found on equidistant grid on $[0.1, 0.65]$ with step size 0.025. Parameter setting $n = 100$, $p = 200$, $p_{\text{eval}} = 120$.

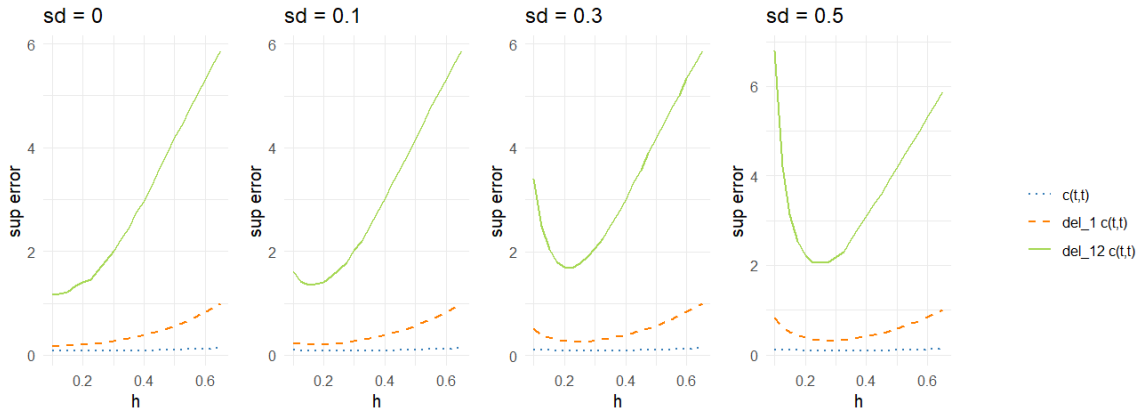

Figure 11: Supremum error for covariance kernel and derivatives for standard deviation of the error. Parameter setting $n = 200$, $p = 120$, $p_{\text{eval}} = 120$.

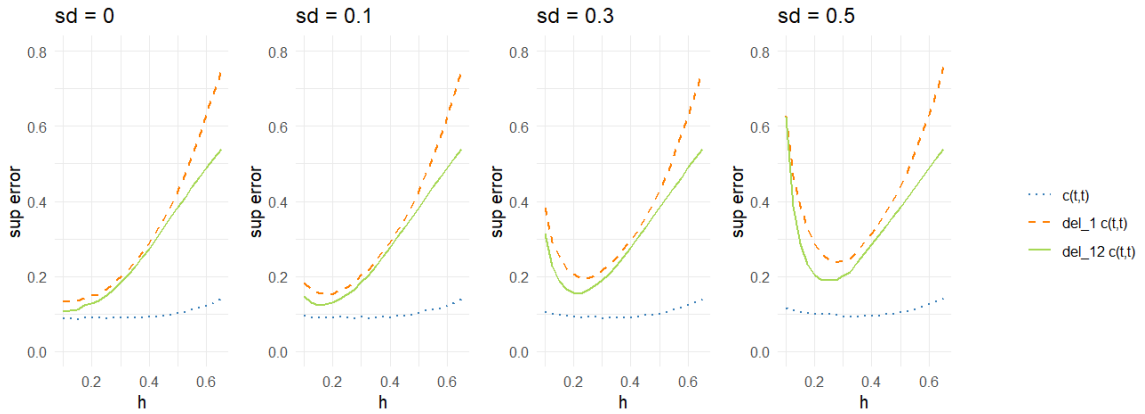

Figure 12: Supremum error (scaled by maximal value of true function) for covariance kernel and derivatives for standard deviation of the error. Parameter setting $n = 200$, $p = 120$, $p_{\text{eval}} = 120$.

σ	$C(t,t)$	$\partial_1 C(t,t)$	$\partial_{12} C(t,t)$
0	0.150	0.100	0.100
0.1	0.325	0.200	0.150
0.3	0.300	0.225	0.225
0.5	0.300	0.275	0.275

Table 8: Optimal bandwidths for different standard deviations of the error, $\sigma \in \{0, 0.1, 0.3, 0.5\}$. Found on equidistant grid on $[0.1, 0.65]$ with step size 0.025. Parameter setting $n = 200$, $p = 120$, $p_{\text{eval}} = 120$.

D. Additional Material for Fairness: Two More Fairness Algorithms and Comparison

We present two more algorithms with which fair confidence bands can be constructed. All three, Algorithm 2 from Subsection 5.2 as well as Algorithm 3 and Algorithm 4, do resemble each other and only differ in the way they choose the correction location t_0. While Algorithm 2 only requires the Kac-Rice formula from Theorem 2.9, the other two require the more general version in Theorem 2.17. The one originally introduced by Liebl and Reimherr (2023) is Algorithm 4 where correction locations are chosen in an alternating way at the interval bounds of the partition.

Algorithm 3: Selecting the fair critical value function $u^*_{\alpha/2}$ for correction location always being at right interval endpoint

---

Input: Provide the roughness parameter function $\tau(t)$, $t \in [0,1]$, set the significance level $\alpha \in (0,1)$ and the number of fairness intervals $p \in \mathbb{N}$.

Initialize the partition $0 = a_0 < \cdots < a_p = 1$ and the critical value function $u_p(t)$ as a piecewise linear function

$$u_p(t) := c_1 + c_2(t - a_1)_+ + \cdots + c_p(t - a_{p-1})_+, \quad t \in [0,1],$$

where $(x)_+ = \max(0, x)$ in order to model $u^*_{\alpha/2}$.

for $j = 1$ **do**

It is $u_p(t) = c_1$ and $u_p'(t) = 0$ for $t \in [0, a_1]$ (constant critical value in first interval). We determine c_1 by solving

$$\mathbb{P}(X(a_1) \geq u(a_1)) + \int_0^{a_1} \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\Big(-\frac{1}{2}c_1^2 \Big)\, dt = \frac{\alpha}{2}(a_1 - 0).$$

for $j = 2, \ldots, p$ **do**

For known $c_1, \ldots, c_{j-1}$, $u_p(t)$ and $u_p'(t)$ for $t \in [a_{j-1}, a_j]$ only depend on c_j the following way

$$u_p(t) = u_p(t, c_j) = c_1 + \sum_{l=1}^{j} c_l (t - a_l)_+ \quad \text{and} \quad u_p'(t) = \sum_{l=1}^{j} c_l.$$

Determine c_j by solving

$$\begin{aligned} &\mathbb{P}(X(a_j) \geq u(a_j)) + \int_{a_{j-1}}^{a_j} \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\Big(-\frac{1}{2}\Big[u_p(t,c_j)^2 + \frac{u_p'(t,c_j)}{\tau(t)^2}\Big]\Big)\, dt \\ &+ \int_{a_{j-1}}^{a_j} \int_0^{\infty} \frac{u_p'(t,c_j)}{2\pi\tau(t)} M_{\mathcal{V}}'\Big(-\frac{1}{2}\Big[u_p(t,c_j)^2 + \frac{(y - u_p'(t,c_j))^2}{\tau(t)^2}\Big]\Big)\, dy\, dt \quad = \frac{\alpha}{2}(a_j - a_{j-1}). \end{aligned}$$

Output: Fair critical value function $u^*_{\alpha/2} = u_p$.

Algorithm 4: Selecting the fair critical value function $u^*_{\alpha/2}$ for correction location at alternating interval endpoints

Input: Provide the roughness parameter function $\tau(t)$, $t \in [0,1]$, set the significance level $\alpha \in (0,1)$ and the number of fairness intervals $p \in \mathbb{N}$.

Initialize the partition $0 = a_0 < \cdots < a_p = 1$ and the critical value function $u_p(t)$ as a piecewise linear function

$$u_p(t) := c_1 + c_2(t - a_1)_+ + \cdots + c_p(t - a_{p-1})_+, \quad t \in [0,1],$$

where $(x)_+ = \max(0, x)$ in order to model $u^*_{\alpha/2}$.

for $j = 1$ **do**

It is $u_p(t) = c_1$ and $u_p'(t) = 0$ for $t \in [0, a_1]$ (constant critical value in first interval). We determine c_1 by solving

$$\mathbb{P}(X(a_1) \geq u(a_1)) + \int_0^{a_1} \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\Big(-\frac{1}{2}c_1^2 \Big)\, dt = \frac{\alpha}{2}(a_1 - 0).$$

for $j = 2, \ldots, p$ **do**

For known $c_1, \ldots, c_{j-1}$, $u_p(t)$ and $u_p'(t)$ for $t \in [a_{j-1}, a_j]$ only depend on c_j the following way

$$u_p(t) = u_p(t, c_j) = c_1 + \sum_{l=1}^{j} c_l (t - a_l)_+ \quad \text{and} \quad u_p'(t) = \sum_{l=1}^{j} c_l.$$

if *j is even* **then**

Determine c_j by solving

$$\begin{aligned} &\mathbb{P}(X(a_{j-1}) \geq u(a_{j-1})) + \int_{a_{j-1}}^{a_j} \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\Big(-\frac{1}{2}\Big[u_p(t,c_j)^2 + \frac{u_p'(t,c_j)}{\tau(t)^2}\Big]\Big)\, dt \\ &\quad - \int_{a_{j-1}}^{a_j} \int_0^{\infty} \frac{u_p'(t,c_j)}{2\pi\tau(t)} M_{\mathcal{V}}'\Big(-\frac{1}{2}\Big[u_p(t,c_j)^2 + \frac{(y + u_p'(t,c_j))^2}{\tau(t)^2}\Big]\Big)\, dy\, dt \quad = \frac{\alpha}{2}(a_j - a_{j-1}) \end{aligned}$$

if *j is odd* **then**

Determine c_j by solving

$$\begin{aligned} &\mathbb{P}(X(a_j) \geq u(a_j)) + \int_{a_{j-1}}^{a_j} \frac{\tau(t)}{2\pi} M_{\mathcal{V}}\Big(-\frac{1}{2}\Big[u_p(t,c_j)^2 + \frac{u_p'(t,c_j)}{\tau(t)^2}\Big]\Big)\, dt \\ &\quad + \int_{a_{j-1}}^{a_j} \int_0^{\infty} \frac{u_p'(t,c_j)}{2\pi\tau(t)} M_{\mathcal{V}}'\Big(-\frac{1}{2}\Big[u_p(t,c_j)^2 + \frac{(y - u_p'(t,c_j))^2}{\tau(t)^2}\Big]\Big)\, dy\, dt \quad = \frac{\alpha}{2}(a_j - a_{j-1}). \end{aligned}$$

Output: Fair critical value function $u^*_{\alpha/2} = u_p$.

Generalized Expected Euler Characteristic Inequality

The following theorem is a generalized version of Theorem 3.1 but with arbitrary correction location $t_0 \in [0,1]$. It shows that $\mathbb{E}[\varphi_{u,X}(t_0)]$ is also an upper bound for the excursion probability of a process X above u. This aspect is particularly important for Algorithms 3 and 4, as it confirms the validity of the confidence bands generated by these algorithms.

Theorem D.1 (Generalized Expected Euler Characteristic Inequality). Let X be a stochastic process and u be a critical value function being defined on $[0,1]$ and satisfying As-

sumption 2.1 and 2.6, then

$$\mathbb{P}\big(\exists\, t \in [0,1] : X(t) \geq u(t)\big) \leq \mathbb{E}\big[\varphi_{u,X}(t_0)\big], \quad t_0 \in [0,1],$$

where $\varphi_{u,X}(t_0)$ is the Euler characteristic as in Definition 2.8.

Proof. To formulate the necessary equality of sets, we define

$$\begin{aligned}
A &= \{X(t_0) \geq u(t_0)\},\\
B &= \{\exists\, t \in [0,t_0) : X(t) > u(t)\},\\
C &= \{\exists\, t \in [0,t_0) : X(t) = u(t), X'(t) > u'(t)\},\\
D &= \{\exists\, t \in [0,t_0) : X(t) = u(t), X'(t) < u'(t)\},
\end{aligned}$$

and B', C', D' respectively over the interval $t \in (t_0, 1]$. Observe that

$$C \subseteq D \cup A \quad \text{and} \quad D' \subseteq C' \cup A \tag{D.1}$$

$$\text{and that} \quad B \subseteq D \cup A \quad \text{and} \quad B' \subseteq C' \cup A. \tag{D.2}$$

We can justify this because, according to Assumption 2.1, the paths of the process $X(t)$ are almost surely continuous. Inclusions (D.1) hold since a continuous path that up-crosses before time t_0 is either greater than $u(t_0)$ at location t_0 or has another down-crossing before t_0 and analog for $(t_0, 1]$. Inclusions (D.2) hold since a continuous path that is above u somewhere before t_0 is either also above $u(t_0)$ at t_0 or has another down-crossing before t_0 and analog for $(t_0, 1]$. Thus, the following equality holds under Assumption 2.6

$$\begin{aligned}
&\{\exists\, t \in [0,1] : X(t) \geq u(t)\}\\
&= \{X(t_0) \geq u(t_0)\} \cup \{\exists\, t \in [0,t_0) : X(t) > u(t)\} \cup \{\exists\, t \in (t_0,1] : X(t) > u(t)\}\\
&\qquad \cup \{\exists\, t \in [0,t_0) : X(t) = u(t)\} \cup \{\exists\, t \in (t_0,1] : X(t) = u(t)\}\\
&\overset{\text{A.2.6}}{=} A \cup B \cup B' \cup C \cup C' \cup D \cup D'\\
&\overset{\text{(D.1)(D.2)}}{=} A \cup D \cup C'\\
&= \{X(t_0) \geq u(t_0)\} \cup \{N_{u,X}^{-1}[0,t_0) \geq 1\} \cup \{N_{u,X}(t_0,1] \geq 1\}.
\end{aligned}$$

The remainder of the inequality is clear using the same arguments as in the paper Liebl and Reimherr (2023) which are the Boole's inequality (union bound) and Markov's inequality and because $\mathbb{P}(N_{u,X}[t_0,t_0] \geq 1) = 0$ and $\mathbb{P}(N_{u,X}^{-1}[t_0,t_0] \geq 1) = 0$. Hence

$$\begin{aligned}
\mathbb{P}(\exists\, t \in [0,1] : X(t) \geq u(t)) = &\quad \mathbb{P}\big(\{X(t_0) \geq u(t_0)\} \cup \{N_{u,X}^{-1}[0,t_0) \geq 1\} \cup \{N_{u,X}(t_0,1] \geq 1\}\big)\\
\overset{\text{Boole}}{\leq} &\quad \mathbb{P}\big(X(t_0) \geq u(t_0)\big) + \mathbb{P}\big(N_{u,X}^{-1}[0,t_0] \geq 1\big) + \mathbb{P}\big(N_{u,X}[t_0,1] \geq 1\big)\\
\overset{\text{Markov}}{\leq} &\quad \mathbb{P}\big(X(t_0) \geq u(t_0)\big) + \mathbb{E}\big[N_{u,X}^{-1}[0,t_0]\big] + \mathbb{E}\big[N_{u,X}[t_0,1]\big]\\
= &\quad \mathbb{E}\big[\varphi_{u,X}(t_0)\big].
\end{aligned}$$

The last equation is Definition 2.16 of the Euler characteristic. □

Price of Fairness of Algorithms 3 and 4

Proposition D.2 (Price of Fairness). The expected Euler characteristic inequality as in Theorem D.1 when using the fair critical value function $u^*_{\alpha/2}$ is given by

$$\mathbb{P}\big(\exists\, t \in [0,1] : X(t) \geq u^*_{\alpha/2}(t)\big) \leq \frac{\alpha}{2} - \varsigma(u^*_{\alpha/2}),$$

where for Algorithm 3, it is

$$\varsigma(u^*_{\alpha/2}) = \begin{cases} 0, & \text{if } 1 = p, \\ \sum_{j=1}^{p} \mathbb{P}(X(a_j) \geq u^*_{\alpha/2}(a_j)) & \text{if } p \geq 2. \end{cases} \tag{D.3}$$

and for Algorithm 4, it is

$$\varsigma(u^*_{\alpha/2}) = \begin{cases} 0, & \text{if } 1 = p, \\ \mathbb{P}(X(a_1) \geq u^*_{\alpha/2}(a_1)) + 2\sum_{\substack{2 \leq j < p \\ j \text{ odd}}} \mathbb{P}(X(a_j) \geq u^*_{\alpha/2}(a_j)) & \text{if } p \geq 2 \text{ and p even}, \\ \mathbb{P}(X(a_1) \geq u^*_{\alpha/2}(a_1)) + 2\sum_{\substack{2 \leq j < p \\ j \text{ odd}}} \mathbb{P}(X(a_j) \geq u^*_{\alpha/2}(a_j)) & \\ +\mathbb{P}(X(a_p) \geq u^*_{\alpha/2}(a_p)) & \text{if } p \geq 2 \text{ and p odd}. \end{cases} \tag{D.4}$$

Formula (D.4) is a corrected version of the price of fairness in Proposition 3.1 in Liebl and Reimherr (2023).

Proof. This proof works analogous to the one of Theorem 5.5. Again, similar to Remark A.7, we add a subscript $[a, b]$ indicating the interval the Euler characteristic $\varphi_{u,X,[a,b]}(t_0)$ is taken on. Note that, in expectation, we can not only sum up $\mathbb{E}\big[N_{u,X}[a, b]\big] + \mathbb{E}\big[N_{u,X}[b, c]\big] = \mathbb{E}\big[N_{u,X}[a, c]\big]$. but also $\mathbb{E}\big[N^{-1}_{u,X}[a, b]\big] = \mathbb{E}\big[N_{u,X}[a, b]\big]$ due to symmetry of X.
Now, on each subinterval, Algorithm 3 solves

$$\mathbb{E}[\varphi_{u,X,[a_{j-1},a_j]}(a_j)] = \frac{\alpha}{2}(a_j - a_{j-1}), \qquad j = 1, \ldots, p.$$

Summing this up for $j = 1, ..., p$ yields $\alpha/2$ for the right side and for the left side

$$\begin{aligned} \sum_{j=1}^{p} \mathbb{E}\big[\varphi_{u,X,[a_{j-1},a_j]}(a_j)\big] &= \sum_{j=1}^{p} \mathbb{P}(X(a_j) \geq u(a_j)) + \mathbb{E}\big[N^{-1}_{u,X}[a_{j-1}, a_j]\big] \\ &= \mathbb{P}(X(a_1) \geq u(a_1)) + \mathbb{E}\big[N^{-1}_{u,X}[0, a_1]\big] + \mathbb{E}\big[N_{u,X}[a_1, 1]\big] + \sum_{j=1}^{p} \mathbb{P}(X(a_j) \geq u(a_j)) \\ &= \mathbb{E}\big[\varphi_{u,X,[0,1]}(a_1)\big] + \varsigma(u), \end{aligned}$$

where ς is the price of fairness as in Equation (D.3). On the contrary, Algorithm 4 solves

$$\begin{aligned} \mathbb{E}[\varphi_{u,X,[a_{j-1},a_j]}(a_{j-1})] &= \frac{\alpha}{2}(a_j - a_{j-1}), \quad \text{for even } j, \\ \mathbb{E}[\varphi_{u,X,[a_{j-1},a_j]}(a_j)] &= \frac{\alpha}{2}(a_j - a_{j-1}), \quad \text{for odd } j. \end{aligned} \tag{D.5}$$

Summing this up for $j = 1, ..., p$ yields again $\alpha/2$ for the right side and for the left side

$$\begin{aligned} &\sum_{j=1}^{p} \mathbb{1}_{\{j \text{ odd}\}} \left(\mathbb{E}\big[\varphi_{u,X,[a_{j-1},a_j]}(a_j)\big] + \mathbb{E}\big[\varphi_{u,X,[a_j,a_{j+1}]}(a_j)\big] \right) \\ &= \sum_{j=1}^{p} \mathbb{1}_{\{j \text{ odd}\}} \left(2\mathbb{P}(X(a_j) \geq u(a_j)) + \mathbb{E}\big[N^{-1}_{u,X}[a_{j-1}, a_j]\big] + \mathbb{E}\big[N_{u,X}[a_j, a_{j+1}]\big] \right) \\ &= \mathbb{E}\big[N^{-1}_{u,X}[0, a_1]\big] + \mathbb{E}\big[N_{u,X}[a_1, 1]\big] + \sum_{j=1}^{p} \mathbb{1}_{\{j \text{ odd}\}} 2\mathbb{P}(X(a_j) \geq u(a_j)) \\ &= \mathbb{E}\big[\varphi_{u,X,[0,1]}(a_1)\big] + \varsigma(u), \end{aligned}$$

where ς is the price of fairness as in Equation (D.4). Hence, with Algorithms 3 and 4, we obtain a critical value function $u^*_{\alpha/2}$ that satisfies

$$\mathbb{E}\big[\varphi_{u^*_{\alpha/2},X}(a_1)\big] + \varsigma(u^*_{\alpha/2}) = \frac{\alpha}{2}.$$

Since already from Theorem D.1, we have

$$\mathbb{P}(\exists\, t \in [0,1] : X(t) \geq u^*_{\alpha/2}(t)) \leq \mathbb{E}[\varphi_{u^*_{\alpha/2},X}(a_1)],$$

we actually obtain a critical value function $u^*_{\alpha/2}$ that satisfies

$$\mathbb{P}(\exists\, t \in [0,1] : X(t) \geq u^*_{\alpha/2}(t)) \leq \frac{\alpha}{2} - \varsigma(u^*_{\alpha/2}), \quad \text{where } \varsigma(u^*_{\alpha/2}) > 0.$$

□

Comparison of Price of Fairness between Algorithms

In Proposition 5.5 and Proposition D.2, we observe that the price of fairness varies depending on the algorithm, i.e., depending on the choice of the correction location. A brief simulation study was conducted to investigate the extent of the discrepancy in the prices of fairness.

We stay with the covariance scenarios as outlined in Section 6 and assume a mean of zero. A sample $n = 100$ processes is generated on a grid of $p = 120$ design points from which we calculate the roughness function τ using the *ffscb* approach (as in Subsection 4.2). Afterwards, fair confidence bands are constructed over a partition of p fairness intervals using all three Algorithms 2, 3 and 4. Note that the Kac-Rice formula for t-distributed processes X, as in Corollary 2.13, was employed since the covariance function has been estimated. In order to calculate the price of fairness, we use the respective formula for each algorithm, given in Proposition 5.5 and n D.2. The prices of fairness are averaged over $N = 1000$ repetitions. This yields the results displayed in Tables 9 and 10 for $p = 4$ and $p = 8$, respectively.

	left (2)	right (3)	alternating (4)
Cov1	0.000801	0.000801	0.000811
Cov2	0.000800	0.000800	0.000810
Cov3	0.000801	0.000800	0.000809

Table 9: Comparison of price of fairness ς for different algorithms and for $p = 4$ fairness intervals.

	left (2)	right (3)	alternating (4)
Cov1	0.001550	0.001552	0.001606
Cov2	0.001550	0.001550	0.001602
Cov3	0.001550	0.001550	0.001607

Table 10: Comparison of price of fairness ς for different algorithms and for $p = 8$ fairness intervals.

It is evident that the prices of fairness are likely to be identical for all algorithms. In both cases, $p = 4$ in Table 9 and $p = 8$ in Table 10, the prices of fairness coincide for each covariance scenario regardless of the algorithm.

Declaration

Herewith I, Lisa Drescher, matriculation number 3106903, declare that this master's thesis with the title "Confidence Bands in Functional Data Analysis based on the Kac-Rice Formula" is a result of my independent work. All sources and auxiliary materials used by me in this thesis are cited completely. This paper was not previously presented to another examination board and has not been published.

Marburg, May 10, 2024

Lisa Drescher