Computational Statistics
https://doi.org/10.1007/s00180-020-01043-6

ORIGINAL PAPER

Simultaneous confidence bands for comparing variance functions of two samples based on deterministic designs

Chen Zhong[1] · **Lijian Yang**[1]

Received: 8 August 2020 / Accepted: 22 October 2020

Abstract

Asymptotically correct simultaneous confidence bands (SCBs) are proposed in both multiplicative and additive form to compare variance functions of two samples in the nonparametric regression model based on deterministic designs. The multiplicative SCB is based on two-step estimation of ratio of the variance functions, which is as efficient, up to order $n^{-1/2}$, as an infeasible estimator if the two mean functions are known a priori. The additive SCB, which is the log transform of the multiplicative SCB, is location and scale invariant in the sense that the width of SCB is free of the unknown mean and variance functions of both samples. Simulation experiments provide strong evidence that corroborates the asymptotic theory. The proposed SCBs are used to analyze several strata pressure data sets from the Bullianta Coal Mine in Erdos City, Inner Mongolia, China.

Keywords Brownian motion · B-spline · Kernel · Oracle efficiency · Strata pressure · Variance ratio

1 Introduction

Nonparametric simultaneous confidence band (SCB) is a useful tool for statistical inference about the global properties of an entire unknown curve or function. It was first constructed in Bickel and Rosenblatt (1973) for a kernel density function. Then nonparametric SCB was soon extended to regression function, see Johnston (1982), Härdle (1989), Härdle and Marron (1991), Eubank and Speckman (1993), Xia (1998), and Claeskens and Van Keilegom (2003) for early works about SCB. SCB not only is a theoretically beautiful construct, but also has wide applications in many areas such as sample survey and functional data analysis, see Zhao and Wu (2008), Ma et al.

✉ Lijian Yang
yanglijian@tsinghua.edu.cn

1 Center for Statistical Science and Department of Industrial Engineering, Tsinghua University, Beijing 100084, China

Published online: 31 October 2020

(2012), Cao et al. (2012, 2016), Song et al. (2014), Wang et al. (2014), Zheng et al. (2014, 2016), Gu et al. (2014), Cai and Yang (2015), Gu and Yang (2015), Wang et al. (2016), Zhang and Yang (2018), and Cai et al. (2020) for recent development on nonparametric SCBs.

In the context of nonparametric regression model, adaptive SCB for the regression function was studied in Hall and Titterington (1988). A rather undesirable limitation of adaptive SCBs is their reliance on the assumption of i.i.d. Gaussian errors and heteroscedasticity (constant variance function). Alternatively, Eubank and Speckman (1993) obtained the SCB for the mean function based on kernel smoothing without Gaussianity assumption on errors; however, it was also under the restrictive assumption of homoscedasticity, and the mean function being periodic. Wang (2012) constructed a spline SCB for nonparametric mean function based on deterministic designs and strongly mixing dependent errors, but the SCB is asymptotically conservative rather than correct. For variance function estimation, Brown and Levine (2007) and Levine (2006) proposed difference-based kernel estimators and an approach of bandwidth selection, but without SCB. Song and Yang (2009) and Cai and Yang (2015) had investigated the SCB for the variance function based on random design, while for deterministic design, Cai et al. (2019) has provided theoretically justified SCBs for both mean and variance functions. All these existing works on SCB concern exclusively one sample problems. The current work extends these to two sample comparison problems.

Testing hypotheses about the difference of two means led to assumptions on the ratio of population variances, see Welch (1938) and James (1951). When the samples were drawn from a given normal bivariate population, Fisher (1924) derived the distribution of log ratio of two sample variance, see also Bose and Mahalanobis (1935), Finney (1938), Scheffé (1942) and Gayen (1950) for distribution theory of the ratio of sample variances. In this work, we propose additive and multiplicative forms of asymptotically correct simultaneous confidence band (SCB) that are independent of mean and variance functions (therefore location and scale invariant), for comparing the variance functions from two independent nonparametric regression.

To be more precise, denote by $\left\{\left(X_{s,i}, Y_{s,i}\right)\right\}_{i=1}^{n_s}$, $s = 1, 2$ the two samples with sample sizes n_s. Often encountered in applications (e.g., the strata pressure data discussed in Sect. 5.2) is the so-called deterministic design nonparametric regression model:

$$Y_{s,i} = m_s\left(\frac{i}{n_s}\right) + \sigma_s\left(\frac{i}{n_s}\right)\varepsilon_{s,i}, \quad i = 1, \ldots, n_s, \quad s = 1, 2 \tag{1}$$

in which the $Y_{s,i}$'s are responses at equally spaced design points i/n_s, $1 \leq i \leq n_s$, and $\left\{\varepsilon_{s,i}\right\}_{i=1}^{n_s}$ are unobserved i.i.d. random errors with $\mathbb{E}\left(\varepsilon_{s,1}\right) = 0$, $\operatorname{var}\left(\varepsilon_{s,1}\right) = 1$. Suppose that the unknown mean and variance functions $m_s(\cdot)$ and $\sigma_s^2(\cdot)$ in model (1) are smooth, Jiang et al. (2020) has established asymptotically correct SCB for the difference $m_1(\cdot) - m_2(\cdot)$ of the two mean functions, under a somewhat surprising assumption that the ratio of two variance functions is a constant: $\sigma_1^2(\cdot)/\sigma_2^2(\cdot) \equiv a^2$. As this thought-provoking assumption itself requires testing, asymptotically correct SCBs are constructed in this paper for the ratio $\sigma_1^2(\cdot)/\sigma_2^2(\cdot)$ as well as its logarithm $\ln\sigma_1^2(\cdot) - \ln\sigma_2^2(\cdot)$. To illustrate the usage of the proposed method, 95% SCB for the variance ratio functions is constructed for several strata pressure data sets collected from the Bulianta Coal Mine located in Erdos City, Inner Mongolia, China. Meanwhile, the SCB for the ratio of two-sample variance functions is used to test the null hypothesis $H_0 : \sigma_1^2(x)/\sigma_2^2(x) \equiv r$ for some constant $r > 0$. Figures 2 and 3 depict the SCB for four pairs of strata pressure data, and the conclusions are weak rejection for the fourth pair with p-value $= 0.023$ and no rejection for the first three pairs with larger p-values.

The remainder of the paper is organized as follows. Section 2 states the main asymptotic theoretical results. Section 3 provides insight into proofs and Sect. 4 presents concrete steps to implement the SCB. Section 5 reports some simulation results and analysis of the strata pressure data. The lemmas and proofs are given in the "Appendix".

2 Main result

In this section the SCB is formulated for the ratio of nonparametric regression variance functions $\sigma_1^2(x)/\sigma_2^2(x)$ in model (1). The variance function $\sigma_s^2(\cdot)$, $s = 1, 2$ measures the heteroscedastic variation of the errors $e_{s,i} = Y_{s,i} - m_s(i/n_s) = \sigma_s(i/n_s)\varepsilon_{s,i}$, $1 \le i \le n_s$ in model (1). Clearly $\mathrm{E}\left(e_{s,i}^2\right) = \sigma_s^2(i/n_s)$, $\mathrm{E}\left(e_{s,i}^4\right) = \sigma_s^4(i/n_s)\mu_{s,4}$, $s = 1, 2$ in which $\mu_{s,4}$ are fourth moments of $\varepsilon_{s,i}$, see Assumption (A2) below. Denote $\sigma_{0,s}^2(x) = \sigma_s^4(x)\left(\mu_{s,4} - 1\right)$ then $\mathrm{var}(e_{s,i}^2) = \mathrm{E}\left(e_{s,i}^4\right) - \left\{\mathrm{E}\left(e_{s,i}^2\right)\right\}^2 = \sigma_{0,s}^2(i/n_s)$.

Following Cai and Yang (2015), if $m_s(x)$ were known by 'oracle', one can begin with smoothing pseudo data sets $\left\{\left(i/n_s, e_{s,i}^2\right)\right\}_{i=1}^{n_s}$ to obtain an infeasible oracle estimator $\tilde{\sigma}_s^2(\cdot)$ for $\sigma_s^2(\cdot)$ $s = 1, 2$. Consequently, a plug-in oracle estimator for $\sigma_1^2(x)/\sigma_2^2(x)$ is $\tilde{\sigma}_1^2(x)/\tilde{\sigma}_2^2(x)$. The oracle estimators are

$$\begin{aligned} \tilde{\sigma}_s^2(x) &= \frac{n_s^{-1}\sum_{i=1}^{n_s} K_h(i/n_s - x)\, e_{s,i}^2}{n_s^{-1}\sum_{i=1}^{n_s} K_h(i/n_s - x)} \\ &= n_s^{-1}\hat{f}_s^{-1}(x)\sum_{i=1}^{n_1} K_h(i/n_s - x)\, e_{s,i}^2, \quad s = 1, 2, \end{aligned} \tag{2}$$

in which $\hat{f}_s(x) = n_s^{-1}\sum_{i=1}^{n_s} K_h(i/n_s - x)$, $s = 1, 2$, $K(u)$ is a kernel function, $h = \max(h_{n_1}, h_{n_2})$, where h_{n_1}, h_{n_2} are sequences of smoothing parameters called bandwidths, and $K_h(u) = h^{-1}K(u/h)$ is the kernel function rescaled by h. However the $\tilde{\sigma}_s^2(x)$'s are infeasible as the errors $\left\{e_{s,i}\right\}_{i=1}^{n}$, $s = 1, 2$ are unobservable. Cai et al. (2019) proposed spline-kernel estimators $\hat{\sigma}_s^2(x)$ to mimic $\tilde{\sigma}_s^2(x)$

$$\hat{\sigma}_s^2(x) = n_s^{-1}\hat{f}_s^{-1}(x)\sum_{i=1}^{n_s} K_h(i/n_s - x)\,\hat{e}_{s,i}^2, \quad s = 1, 2,$$

where $\hat{e}_{s,i} = Y_{s,i} - \hat{m}_{s,p}(i/n_s), s = 1, 2$ and $\hat{m}_{s,p}(x), s = 1, 2$ are the pth order spline estimator for $m_s(x)$ with integer $p > 0$,

$$\hat{m}_{s,p}(x) = \arg \min_{g \in \mathcal{H}_N^{(p-2)}} \sum_{i=1}^{n_s} \left\{ Y_{s,i} - g(i/n) \right\}^2, \tag{3}$$

in which $\mathcal{H}_N^{(p-2)} = \mathcal{H}_N^{(p-2)}[0, 1]$ is the space of spline functions on interval $[0, 1]$ defined below.

Divide the interval $[0, 1]$ into $(N + 1)$ subintervals $J_j = \left[\chi_j, \chi_{j+1}\right), j = 0, 1, 2, \ldots, N$ by equally spaced points $\left\{\chi_j\right\}_{j=1}^{N}$ called interior knots,

$$0 = \chi_0 < \chi_1 < \cdots < \chi_{N+1} = 1, \; \chi_j = j/(N + 1), \; j = 0, 1, \ldots, N + 1.$$

$\mathcal{H}_N^{(p-2)}$ is the space of functions that are polynomials of degree $(p - 1)$ on each J_j with continuous $(p - 2)$th derivative on $[0, 1]$. For instance, $\mathcal{H}_N^{(-1)}$ consists of functions that are constant on each J_j, and $\mathcal{H}_N^{(0)}$ the space of functions that are linear on each J_j and continuous on $[0, 1]$.

For $s = 1, 2$, the estimator $\hat{m}_{s,p}(x)$ in (3) can be expressed as

$$\hat{m}_{s,p}(x) = \sum_{j=1-p}^{N_s} \hat{\lambda}_{s,j,p} B_{j,p}(x),$$

where the vector $\left(\hat{\lambda}_{s,1-p,p}, \ldots, \hat{\lambda}_{s,N,p}\right)^T$ is the solution of the least-squares problem

$$\left(\hat{\lambda}_{s,1-p,p}, \ldots, \hat{\lambda}_{s,N,p}\right)^T = \operatorname{argmin}_{\mathbb{R}^{N_s+p}} \sum_{i=1}^{n_s} \left\{ Y_{s,i} - \sum_{j=1-p}^{N_s} \lambda_{s,j,p} B_{j,p}(x) \right\}^2. \tag{4}$$

Denote by $\psi^{(s)}(x)$ the sth order derivative of a function $\psi(x)$. For $\theta \in (0, 1]$ and integer $p \geq 0$, let $C^{p,\theta}[0, 1]$ be the space of functions with θ−Hölder continuous pth-order derivatives on $[0, 1]$ with seminorm $\|\cdot\|_{p,\theta}$

$$C^{p,\theta}[0, 1] = \left\{ \phi(x) : \|\phi\|_{p,\theta} = \sup_{x \neq x', x, x' \in [0,1]} \frac{\left|\phi^{(p)}(x) - \phi^{(p)}\left(x'\right)\right|}{|x - x'|^{\theta}} < +\infty \right\},$$

and denote by $C^{(p)}[0, 1]$ the space of p-times continuously differentiable functions. For sequences of positive real numbers c_n and d_n, $c_n \ll d_n$ means $c_n/d_n \to 0$ as $n \to \infty$.

Denote $n = \min(n_1, n_2)$, and $\eta_{s,i} = \left(\varepsilon_{s,i}^2 - 1\right)\left(\mu_{s,4} - 1\right)^{-1/2}$, then $\mathbb{E}(\eta_{s,i}) = 0$, $\mathbb{E}(\eta_{s,i}^2) = 1, s = 1, 2$. We need the following assumptions to construct SCBs for $\sigma_1^2(x)/\sigma_2^2(x)$.

(A1) *The functions* $m_s(\cdot) \in C^p[0,1]$, $s = 1, 2$ *for integer* $p > 1$.

(A2) *The errors* $\varepsilon_{s,i}$, $s = 1, 2$ *satisfy* $\mathbb{E}(\varepsilon_{s,i}) = 0$, $\mathbb{E}(\varepsilon_{s,i}^2) = 1$, $\mathbb{E}(\varepsilon_{s,i}^4) = \mu_{s,4} < \infty$ *and* $\sigma_s^2(\cdot) \in C^{p_0-1,\theta_0}[0,1]$ *for integer* $p_0 > 1$, $\theta_0 \in (0,1]$ *with* $0 < c_\sigma \leq \sigma_s^2(x) \leq C_\sigma < +\infty$ *for any* $x \in [0,1]$.

(A3) *There exist constants* $c, C \in (0,\infty)$ *such that* $0 < c \leq n_1/n_2 \leq C < \infty$ *as* $n \to \infty$.

(A4) *There exist* $\beta_s' \in (0, 1/2 - 1/(4\theta_0 + 4p_0 - 2))$, $C_s' \in (0,+\infty)$, $\gamma_s' \in (1,+\infty)$ *and i.i.d.* $N(0,1)$ *variables* $\left\{Z_{s,in_s}'\right\}_{i=1}^{n_s}$, $s = 1, 2$ *such that*

$$\mathbb{P}\left\{\max_{1\leq l\leq n_s}\left|\sum_{i=1}^{l}\eta_{s,i} - \sum_{i=1}^{l}Z_{s,in_s}'\right| > n_s^{\beta_s'}\right\} < C_s n_s^{-\gamma_s'}.$$

(A5) *There exist* $C_s \in (0,+\infty)$, $\gamma_s \in (1,+\infty)$, $\beta_s \in (0,b]$, *and i.i.d.* $N(0,1)$ *variables* $\left\{Z_{s,in_s}\right\}_{i=1}^{n}$, $s = 1, 2$ *such that*

$$P\left\{\max_{1\leq l\leq n_s}\left|\sum_{i=1}^{l}\varepsilon_{s,i} - \sum_{i=1}^{l}Z_{s,in_s}\right| > n^{\beta_s}\right\} < C_s n^{-\gamma_s},$$

where $b = \min\{1 - 3/2(2p+1) - t, 1 - 5/2(2p+1) - 5t/2(2p+3), 1/2 - 1/(4\theta_0 + 4p_0 - 2)\}$.

(A6) *The kernel function* $K \in C^{(1)}(\mathbb{R})$, *is of order* p_0, *and is supported on* $[-1,1]$.

(A7) *The bandwidths* h_{n_s}, $s = 1, 2$, *satisfy* $\log h_{n_s}/(-\log n_s) \to t > 0$ *as* $n \to \infty$ *and*

$$\max\left\{n_s^{-1/2}\log^{1/2}n_s, n_s^{2\beta_s'-1}\log n_s, n_s^{-2(p-1)/(2p+1)}\right\} \ll h_{n_s}$$

$$\ll (n_s \log n_s)^{-1/(2\theta_0+2p_0-1)}.$$

Hence $1/(2\theta_0 + 2p_0 - 1) \leq t \leq \min\left\{1/2, 1 - 2\max\left\{\beta_1', \beta_2'\right\}, 2(p-1)/(2p+1)\right\}$.

(A8) *The number of interior knots* N_s *satisfies* $\log N_s/\log n_s \to \tau$ for some $\tau_s > 0$, $s = 1, 2$ *and*

$$\max\left\{n_s^{1/4p}, h_{n_s}^{-1/(p-1)}n_s^{(\beta_s-1/2)/(p-1)}, h_{n_s}^{-1/2(p-1)}\log^{1/2(p-1)}n_s\right\} \ll N_s$$

$$\ll \min\left\{h_{n_s}{}^{2/3}n_s^{2(1-\beta_s)/3}, n_s^{2(1-\beta_s)/5}, n_s^{1/3}h_{n_s}^{1/3}\log^{-1/3}n_s\right\},$$

Consequently, $\max\{1/4p, (2t + 2\beta_s - 1)/2(p-1), t/2(p-1)\} \leq \tau_s \leq \min\{2(1-\beta_s)/3 - 2t/3, 2(1-\beta_s)/5.1/3 - t/3\}$.

Assumption (A1) is a general condition for spline regression of the mean function in model (1). Assumption (A2) is adopted from Härdle (1989). Assumption (A3) requires only that the ratio of sample sizes is comparable, and Assumption (A6) is standard for kernel function. Assumption (A7) is a general condition on the selection of bandwidth h_s to ensure the asymptotic distribution of $\ln \frac{\tilde{\sigma}_1^2(x)}{\tilde{\sigma}_2^2(x)}$. Assumption (A8) on the choice of knot number N_s guarantees the oracle efficiency in Theorem 1. According to Cai et al. (2019), (A4) and (A5) are ensured by Assumption (A4').

(A4') *There exists* $\eta_s' > 2/\beta_s - 2$, $\beta_s \in (0, b\,]$ *as in (A5) such that* $\mathrm{E}\left|\varepsilon_{s,1}\right|^{4+2\eta_s'} < +\infty$.

In order to construct SCB for $\sigma_1^2(x)/\sigma_2^2(x)$, one first constructs SCB for $\ln \sigma_1^2(x) - \ln \sigma_2^2(x)$, and then takes exponential transformation to obtain the desired result. From now on, denote $\mathcal{I}_n = [h_n, 1-h_n]$. Proofs of the following Propositions 1 and 2 are in the "Appendix".

Proposition 1 *Under Assumptions (A2)–(A5), as* $n \to \infty$,

$$\mathbb{P}\left[a_h\left\{v_n^{-1}\sup_{x\in\mathcal{I}_n}\left|\ln\frac{\tilde{\sigma}_1^2(x)}{\tilde{\sigma}_2^2(x)} - \ln\frac{\sigma_1^2(x)}{\sigma_2^2(x)}\right| - b_h\right\} \le z\right] \to \exp\{-2\exp(-z)\},\ z \in \mathbb{R},$$

where $a_h = \left\{2\log\left(h^{-1}\right)\right\}^{1/2}$, $b_h = a_h + a_h^{-1}\left\{2^{-1}\log\left(C_K/\left(4\pi^2\right)\right)\right\}$,

$$C_K = \int_{-1}^{1} K^{(1)}(v)^2\,dv \Big/ \int_{-1}^{1} K(v)^2\,dv,$$
$$v_n = h^{-1/2}\left[\left\{n_1^{-1}\left(\mu_{1,4}-1\right) + n_2^{-1}\left(\mu_{2,4}-1\right)\right\}\int_{-1}^{1} K^2(u)\,du\right]^{1/2}. \tag{5}$$

Proposition 2 *Under Assumptions (A1)–(A8), as n* $\to \infty$ *the spline-kernel estimator* $\ln \frac{\hat{\sigma}_1^2(x)}{\hat{\sigma}_2^2(x)}$ *is asymptotically as efficient as the 'infeasible' estimator* $\ln \frac{\tilde{\sigma}_1^2(x)}{\tilde{\sigma}_2^2(x)}$ *in the sense that*

$$\sup_{x\in\mathcal{I}_n}\left|\ln\frac{\hat{\sigma}_1^2(x)}{\hat{\sigma}_2^2(x)} - \ln\frac{\tilde{\sigma}_1^2(x)}{\tilde{\sigma}_2^2(x)}\right| = o_p\left(n^{-1/2}\right).$$

Consequently, since $a_h \sim \log^{1/2} n$, $v_n^{-1} = \mathcal{O}_p\left(n^{1/2}h^{1/2}\right)$,

$$a_h v_n^{-1}\sup_{x\in\mathcal{I}_n}\left|\ln\frac{\hat{\sigma}_1^2(x)}{\hat{\sigma}_2^2(x)} - \ln\frac{\tilde{\sigma}_1^2(x)}{\tilde{\sigma}_2^2(x)}\right| = o_p(1).$$

Propositions 1 and 2, Slutsky's Theorem together imply the following.

Theorem 1 *Under Assumptions (A1)–(A8),as n* $\to \infty$,

$$\mathbb{P}\left[a_h\left\{v_n^{-1}\sup_{x\in\mathcal{I}_n}\left|\ln\frac{\hat{\sigma}_1^2(x)}{\hat{\sigma}_2^2(x)} - \ln\frac{\sigma_1^2(x)}{\sigma_2^2(x)}\right| - b_h\right\} \le z\right] \to \exp\{-2\exp(-z)\},$$

in which a_h, b_h, C_K *and* v_n *given in* (5). *Then, for any* $\alpha \in (0,1)$,

$$\mathbb{P}\left\{\ln\frac{\sigma_1^2(x)}{\sigma_2^2(x)} \in \ln\frac{\hat\sigma_1^2(x)}{\hat\sigma_2^2(x)} \pm v_n\left(a_h^{-1}q_\alpha + b_h\right), \forall x \in \mathcal{I}_n\right\} \to 1-\alpha,$$

where $q_\alpha = -\log\{-1/2\log(1-\alpha)\}$. *Consequently, for any* $\alpha \in (0,1)$,

$$\mathbb{P}\left[\frac{\hat\sigma_1^2(x)}{\hat\sigma_2^2(x)}\exp\left\{-v_n\left(a_h^{-1}q_\alpha + b_h\right)\right\} \le \frac{\sigma_1^2(x)}{\sigma_2^2(x)} \le \frac{\hat\sigma_1^2(x)}{\hat\sigma_2^2(x)}\exp\left\{v_n\left(a_h^{-1}q_\alpha + b_h\right)\right\}, \forall x \in \mathcal{I}_n\right] \to 1-\alpha. \tag{6}$$

Theorem 1 implies that the additive SCB's contracting width is $\left\{n_1^{-1}\left(\mu_{1,4}-1\right) + n_2^{-1}\left(\mu_{2,4}-1\right)\right\}^{1/2} h^{-1/2}\log^{1/2}h^{-1}$, which does not depend on the unknown mean functions $m_s(\cdot)$ and variance functions $\sigma_s^2(\cdot)$, $s = 1, 2$, in stark contrast to the SCB for variance function of one sample in Song and Yang (2009), Cai and Yang (2015), and Cai et al. (2019). In the special case $p = 4$, $p_0 = 2$, $\theta_0 = 1$ as in Subsection 4.1, the implemented order of h satisfying Assumption (A7) is $n_1^{-1/5}\log^{-1/5-\delta_1}n_1$ or $n_2^{-1/5}\log^{-1/5-\delta_1}n_2$ for any $\delta_1 > 0$. Thus, the optimal bandwidth order is undersmoothed by $\log^{-1/5-\delta_1}n_1$ or $\log^{-1/5-\delta_1}n_2$, and the contracting width of the additive SCB is $\left\{n_1^{-1}\left(\mu_{1,4}-1\right) + n_2^{-1}\left(\mu_{2,4}-1\right)\right\}^{1/2} n_1^{1/10}\log^{3/5+0.5\delta_1}n_1$.

3 Error decomposition

Asymptotic SCB for $\ln\sigma_1^2(x) - \ln\sigma_2^2(x)$ is constructed starting with investigating $\sup_{x\in\mathcal{I}_n}\left|\ln\frac{\tilde\sigma_1^2(x)}{\tilde\sigma_2^2(x)} - \ln\frac{\sigma_1^2(x)}{\sigma_2^2(x)}\right|$. Consider that

$$\begin{aligned}
\tilde\sigma_s^2(x) - \sigma_s^2(x) &= n_s^{-1}\hat f_s^{-1}(x)\sum_{i=1}^{n_s}K_h\left(\frac{i}{n_s}-x\right)e_{s,i}^2 - \sigma_s^2(x) \\
&= n_s^{-1}\hat f_s^{-1}(x)\sum_{i=1}^{n_s}K_h\left(\frac{i}{n_s}-x\right)\left\{e_{s,i}^2 - \sigma_s^2(x)\right\} \\
&= n_s^{-1}\hat f_s^{-1}(x)\sum_{i=1}^{n_s}K_h\left(\frac{i}{n_s}-x\right)\left\{\sigma_s^2\left(\frac{i}{n_s}\right)\varepsilon_{s,i}^2 - \sigma_s^2(x)\right\} \\
&= \hat f_s^{-1}(x)\left\{A_{s,n_s}(x) + B_{s,n_s}(x)\right\},
\end{aligned}$$

in which

$$A_{s,n_s}(x) = n_s^{-1}\sum_{i=1}^{n_s}K_h\left(\frac{i}{n_s}-x\right)\left\{\sigma_s^2\left(\frac{i}{n_s}\right) - \sigma_s^2(x)\right\}$$

$$B_{s,n_s}(x) = n_s^{-1} \sum_{i=1}^{n_s} K_h\left(\frac{i}{n_s} - x\right)\left\{\sigma_s^2\left(\frac{i}{n_s}\right)\varepsilon_{s,i}^2 - \sigma_s^2\left(\frac{i}{n_s}\right)\right\}$$

Then the following stochastic processes approximate $B_{s,n_s}(x)$:

$$B_{s,n_s,1}(x) = n_s^{-1} \sum_{i=1}^{n_s} K_h(i/n_s - x)\,\sigma_{0,s}(i/n_s) Z'_{s,in_s}, \tag{7}$$

$$B_{s,n_s,2}(x) = n_s^{-1} \sum_{i=1}^{n_s} K_h(i/n_s - x)\,\sigma_{0,s}(x)\, Z'_{s,in_s}, \tag{8}$$

$$B_{s,n_s,3}(x) = n_s^{-1/2} \int K_h(u - x)\,\sigma_{0,s}(x)\, dW_{s,n_s}(u),\ x \in \mathcal{I}_n \tag{9}$$

where $\left\{Z'_{s,in_s}\right\}_{i=1}^{n_s}$ are i.i.d. $N(0, 1)$ variables satisfying (A4) and $W_{s,n_s}(u)$ is a two-sided Brownian motion on $(-\infty, +\infty)$ satisfying

$$Z'_{s,in_s} = \sqrt{n}\left\{W_{s,n_s}(i/n_s) - W_{s,n_s}((i-1)/n_s)\right\}.$$

Define a Gaussian process

$$\zeta(x) = \frac{n_1^{-1/2}\nu_{1,4}^{1/2} \int K(x-r)\, dW_{1,n_1}(r) - n_2^{-1/2}\nu_{2,4}^{1/2} \int K(x-r)\, dW_{2,n_2}(r)}{\left[\left\{n_1^{-1}\nu_{1,4} + n_2^{-1}\nu_{2,4}\right\} \int_{-1}^{1} K^2(u)\, du\right]^{1/2}},$$
$$x \in \left[1, h^{-1} - 1\right] = \mathcal{I}_n/h, \tag{10}$$

in which $\nu_{1,4} = \mu_{1,4} - 1$ and $\nu_{2,4} = \mu_{2,4} - 1$.

The following result is essential for proving Theorem 1.

Proposition 3 *Under Assumptions (A2), (A6) and (A7), as $n \to \infty$,*

$$P\left[a_h\left\{\sup_{x\in[1,h^{-1}-1]} |\zeta(x)| - b_h\right\} < z\right] \to \exp\{-2\exp(-z)\},\ z \in \mathbb{R},$$

where a_h and b_h are given in (5), and $\zeta(x)$ is defined in (10).

The proof of Proposition 3 is given in the "Appendix".

4 Implementation

In this section, we describe detailed procedures for implementing the SCBs in Theorem 1 based on two-sample data sets $\{(i/n_s), Y_{s,i}\}_{i=1}^{n_s}$ in the model (1). This is used throughout in Sect. 5 for simulations and real data examples.

The default values are $p = 4$, $p_0 = 2$, $\theta_0 = 1$ in (A1) and (A2) when constructing the SCBs for the ratio function $\sigma_1^2(x)/\sigma_2^2(x)$ in model (1) according to Theorem 1. Meanwhile, one chooses a kernel function K and bandwidth h for computing the spline-kernel estimates $\hat{\sigma}_s^2(x)$, and then plugs in these estimates.

We choose the quartic kernel $K(u) = 15\left(1 - u^2\right) I\{|u| \leq 1\}/16$ to satisfy (A6), and the bandwidths $h = \max(h_{1,\mathrm{rot}} \times \log^{-1/5-\delta_1} n_1, h_{2,\mathrm{rot}} \times \log^{-1/5-\delta_1} n_2)$ $(\delta_1 > 0)$ to satisfy (A7), where the rule-of-thumb bandwidth $h_{s,\mathrm{rot}}$, $s = 1, 2$ is from Equation (4.3) of Fan and Gijbels (1996):

$$h_{s,\mathrm{rot}} = \left[\frac{35 \sum_{i=1}^{n_s} \left\{ \hat{e}_{s,i}^2 - \sum_{k=0}^{4} \widehat{a}_k (i/n_s)^k \right\}^2}{n_s \sum_{i=1}^{n_s} \left\{ 2\widehat{a}_2 + 6\widehat{a}_3 (i/n_s) + 12\widehat{a}_4 (i/n_s)^2 \right\}^2} \right]^{1/5}, \tag{11}$$

in which $(\widehat{a}_k)_{k=0}^4 = \operatorname{argmin}_{(a_k)_{k=0}^4 \in \mathbb{R}^5} \sum_{i=1}^{n_s} \left(\hat{e}_{s,i}^2 - \sum_{k=0}^{4} a_k (i/n_s)^k \right)^2$. According to (11), $h_{s,\mathrm{rot}}$ is of order $n_s^{-1/5}$ and $h_{n_s} \propto n_s^{-1/5} \log^{-1/5-\delta_1} n_s$, $s = 1, 2$, satisfying (A7). We have found that $h = \max(h_{1,\mathrm{rot}}, h_{2,\mathrm{rot}}) \log^{-1/2}(n_1 + n_2)/2$ works quite well via extensive simulations; thus, that is what we recommend.

According to Theorem 1 of Xue and Yang (2006), for any $m(x) \in C^p[0, 1]$, $p \geq 2$, the optimal order of knots number N_s for $m_s(x)$ is $n_s^{1/(2p+1)}$, $n_s^{1/9}$, $s = 1, 2$ with $p = 4$. Denote the 'optimal' N_s by $\hat{N}_s^{\mathrm{opt}}$, the minimizer of the BIC defined below over integers in $\left[0.5N_{s,r}, \min\left\{5N_{s,r}, Tb\right\}\right]$, where $N_{s,r} = n_s^{1/9}$ and $Tb = n_s/4 - 1$ to ensure that $\hat{N}_s^{\mathrm{opt}}$ is of order $n_s^{1/9}$ and the total parameters in the least square estimation is less than $n_s/4$. This particular $\hat{N}_s^{\mathrm{opt}}$ satisfies (A8), but is certainly not the only one. Let $\hat{Y}_{s,i} = \hat{m}_{s,p}(i/n)$, $s = 1, 2$ be the predictor of the ith response $Y_{s,i}$ and $q_{s,n} = (4 + N_s)$ represent the number of parameters in (4). The BIC value corresponding to N_s is

$$\mathrm{BIC}(N_s) = \log \mathrm{MSE} + q_{s,n} \log n_s/n_s, \quad \mathrm{MSE} = n_s^{-1} \sum_{i=1}^{n_s} (Y_{s,i} - \hat{Y}_{s,i})^2, s = 1, 2. \tag{12}$$

To estimate the centered fourth moment $\mu_{s,4}$ of $\varepsilon_{s,1}$, $s = 1, 2$, one can use the spline estimators $\hat{m}_{s,p}(\cdot)$ and $\hat{\sigma}_s^2(\cdot)$:

$$\hat{\mu}_{s,4} = n_s^{-1} \sum_{i=1}^{n_s} \left\{ \frac{Y_{s,i} - \hat{m}_{s,p}(i/n)}{\hat{\sigma}_s(i/n)} \right\}^4, s = 1, 2.$$

The asymptotic $100(1-\alpha)\,\%$ SCB for variance ratio function $\sigma_1^2(x)/\sigma_2^2(x)$ is:

$$\left[\frac{\hat{\sigma}_1^2(x)}{\hat{\sigma}_2^2(x)}\exp\left\{-\hat{v}_n\left(a_h^{-1}q_\alpha+b_h\right)\right\},\frac{\hat{\sigma}_1^2(x)}{\hat{\sigma}_2^2(x)}\exp\left\{\hat{v}_n\left(a_h^{-1}q_\alpha+b_h\right)\right\}\right],x\in\mathcal{I}_n, \tag{13}$$

with $\hat{v}_n=h^{-1/2}\left[\left\{n_1^{-1}\left(\hat{\mu}_{1,4}-1\right)+n_2^{-1}\left(\hat{\mu}_{2,4}-1\right)\right\}\int_{-1}^{1}K^2(u)\,du\right]^{1/2}$.

5 Empirical studies

5.1 Monte Carlo examples

To investigate the finite-sample behavior of the proposed SCB in Section 2, the following two cases are examined.
Case 1:

$$\begin{aligned}&m_1(x)=\cos(3\pi x),m_2(x)=2m_1(x),\\&\sigma_1(x)=0.1\sin(2\pi x)+0.2,\sigma_2(x)=2\sigma_1(x).\end{aligned}$$

Case 2:

$$\begin{aligned}&m_1(x)=\cos(3\pi x),m_2(x)=2m_1(x),\\&\sigma_1(x)=0.1\sin(2\pi x)+0.2,\sigma_2(x)=\frac{\exp(x/4)-0.9}{\exp(x/4)+0.9}.\end{aligned}$$

The error ε follows $U\left(-\sqrt{3},\sqrt{3}\right)$, $N(0,1)$ or the standardized t-distribution with freedom 10, $\varepsilon\sim 0.8^{1/2}t_{10}$. The sample sizes are $n_1,n_2=300,600,900$, while for the SCB, the confidence level $1-\alpha=0.95,0.99$. The coverage frequencies by SCB defined in (13) for $\sigma_1^2(x)/\sigma_2^2(x)$ are reported in Table 1; these are relative frequencies in 2000 replications of coverage of the true curve at equally spaced points $\{x_i=h+i(1-2h)/n, i=1,2,\ldots,n=\max(n_1,n_2)\}$ on $\mathcal{I}_n$. In all cases with $\varepsilon\sim U\left(-\sqrt{3},\sqrt{3}\right)$, $\varepsilon\sim N(0,1)$ and $\varepsilon\sim 0.8^{1/2}t_{10}$, the coverage frequencies improve and approach the nominal level as the sample size n_1 and n_2 increases, which supports Theorem 1. One also finds that the coverage frequencies for $\varepsilon\sim U\left(-\sqrt{3},\sqrt{3}\right)$ approach the nomial level best, and the coverage frequencies for $\varepsilon\sim N(0,1)$ approach the nomial level better than $\varepsilon\sim 0.8^{1/2}t_{10}$.

To visualize the SCB for the ratio of variance functions, Figure 1 were depicted based on three cases of sample size in Case 1 with either $\varepsilon\sim N(0,1)$ or $\varepsilon\sim U\left(-\sqrt{3},\sqrt{3}\right)$ and confidence level 95%. Each has center solid line as the true curve, center dashed line as the estimated curve and the upper and lower thick lines the SCB.

Table 1 The coverage frequencies of the SCBs in (6) of Theorem 1 for $\sigma_1^2(x)/\sigma_2^2(x)$ based on 2000 replications with $\varepsilon \sim U\left(-\sqrt{3}, \sqrt{3}\right)$, $\varepsilon \sim N(0,1)$ and $\varepsilon \sim 0.8^{1/2}t_{10}$ respectively

n_1	n_2	$1-\alpha$	$\varepsilon \sim U\left(-\sqrt{3}, \sqrt{3}\right)$, $\varepsilon \sim N(0,1)$, $\varepsilon \sim 0.8^{1/2}t_{10}$ Case 1	Case 2
300	300	0.95	0.943, 0.931, 0.919	0.931, 0.918, 0.882
		0.99	0.995, 0.988, 0.983	0.992, 0.987, 0.986
	600	0.95	0.943, 0.929, 0.920	0.924, 0.917, 0.895
		0.99	0.993, 0.991, 0.990	0.991, 0.981, 0.979
	900	0.95	0.938, 0.930, 0.913	0.920, 0.915, 0.883
		0.99	0.988, 0.987, 0.988	0.991, 0.988, 0.979
600	300	0.95	0.939, 0.933, 0.929	0.930, 0.914, 0.903
		0.99	0.993, 0.994, 0.989	0.988, 0.985, 0.986
	600	0.95	0.951, 0.938, 0.934	0.928, 0.936, 0.910
		0.99	0.995, 0.995, 0.995	0.989, 0.992, 0.993
	900	0.95	0.948, 0.948, 0.936	0.927, 0.937, 0.913
		0.99	0.999, 0.997, 0.990	0.989, 0.990, 0.990
900	300	0.95	0.940, 0.933, 0.916	0.917, 0.925, 0.894
		0.99	0.990, 0.988, 0.988	0.983, 0.988, 0.985
	600	0.95	0.958, 0.956, 0.941	0.931, 0.941, 0.918
		0.99	0.996, 0.994, 0.991	0.992, 0.995, 0.986
	900	0.95	0.956, 0.951, 0.952	0.922, 0.941, 0.929
		0.99	0.994, 0.995, 0.996	0.992, 0.992, 0.993

As expected, the SCBs for greater sample size are thinner and fit better than those for smaller sample size.

5.2 Data examples

We have applied the two-sample SCB to data sets obtained from the research group at China University of Mining and Technology headed by Professor Jiang Yaodong. The data consists of strata pressure at the Bulianta Coal Mine located in Erdos City, Inner Mongolia, China recorded in May 2013. Strata pressure patterns, such as the range and pressure periodicity in front of working face, are useful information for improving the safety and accuracy of underground mining. For instance, accidents caused by sudden increase of strata pressure are preventable by appropriate roof support design; see Ju and Xu (2013) and Qian et al. (2010).

Measured in units of KN/m^2, strata pressure is the vertical stress on the coal seam roof in front of the working face, a working face is the underground location where coal is peeled from the coal wall mechanically by miners. Data is collected at a record distance of 0.80 m by pressure sensors placed on top of the hydraulic support in front of the working face. During the mining process, the hydraulic support moves forward at a pace of 0.80 m in the propulsion range from 295.5 to 705.1 m, so the sample size

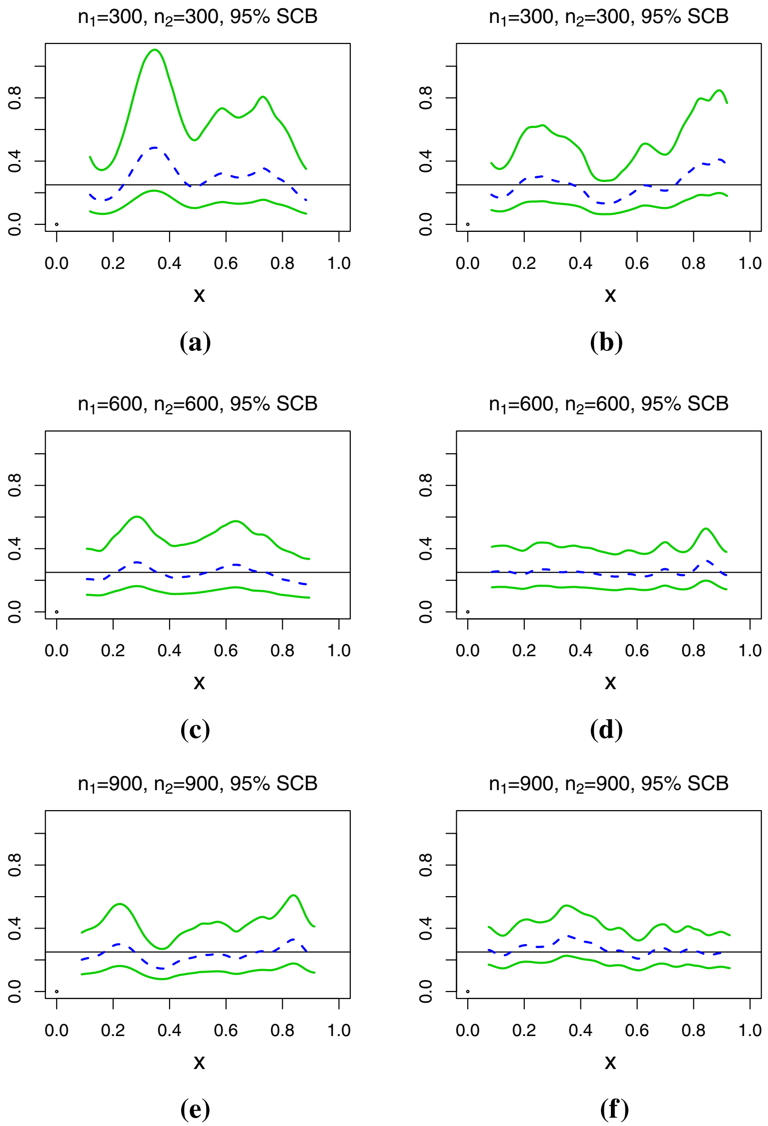

Fig. 1 Plots of 95% SCB (thick) for $\sigma_1^2(x)/\sigma_2^2(x)$ (solid) and the estimator $\hat{\sigma}_1^2(x)/\hat{\sigma}_2^2(x)$ (dashed) in Case 1, with **a** $n_1 = n_2 = 300, \varepsilon \sim N(0,1)$; **b** $n_1 = n_2 = 300, \varepsilon \sim U\left(-\sqrt{3}, \sqrt{3}\right)$; **c** $n_1 = n_2 = 600, \varepsilon \sim N(0,1)$; **d** $n_1 = n_2 = 600, \varepsilon \sim U\left(-\sqrt{3}, \sqrt{3}\right)$; **e** $n_1 = n_2 = 900, \varepsilon \sim N(0,1)$; **f** $n_1 = n_2 = 900, \varepsilon \sim U\left(-\sqrt{3}, \sqrt{3}\right)$

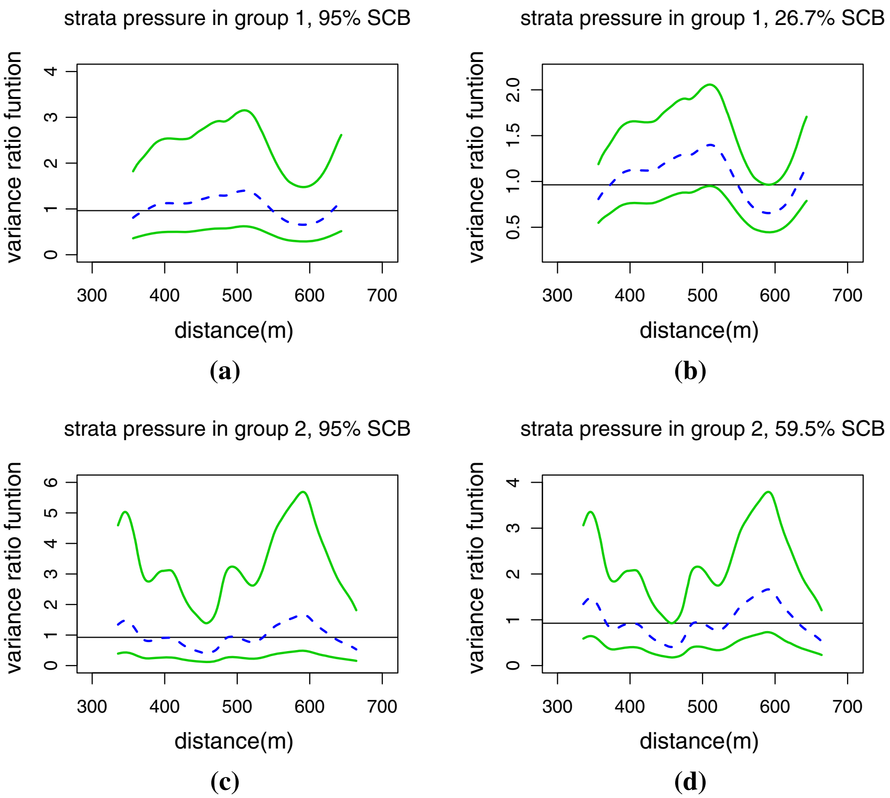

Fig. 2 Plots of the null hypothesis curve of $\hat{r} = n_1^{-1}\sum_{i=1}^{n_1}\hat{e}_{1,i}^2/n_2^{-1}\sum_{i=1}^{n_2}\hat{e}_{2,i}^2$ (solid), SCB (thick) for $\sigma_1^2(x)/\sigma_2^2(x)$ and the spline-kernel estimator $\hat{\sigma}_1^2(x)/\hat{\sigma}_2^2(x)$ (dashed), with **a** 95% SCB for pair 1; **b** lowest simultaneous confidence band containing null hypothesis for pair 1; **c** 95% SCB for pair 2; **d** lowest simultaneous confidence band containing null hypothesis for pair 2

is $n = 1 + (705.1 - 295.5)/0.8 = 513$, and the propulsion range is standardized to interval [0, 1].

The strata pressure data are recorded at 28 sites. Out of the 28 sites, 8 sites are randomly selected, and data sets of sample size 513 from these 8 sites are then randomly divided into 4 pairs for variance function comparison. Figures 2 and 3 show the SCBs (thick lines) computed according to (13) for the function $\sigma_1^2(x)/\sigma_2^2(x)$, and spline-kernel estimate $\hat{\sigma}_1^2(x)/\hat{\sigma}_2^2(x)$ (dashed line). Engineers are interested in whether two different sites have comparable variance functions. One therefore proposes the null hypothesis $H_0: \sigma_1^2(x)/\sigma_2^2(x) \equiv r$ to be tested by the SCB for the ratio of variance functions $\sigma_1^2(\cdot)/\sigma_2^2(\cdot)$. For the four pairs, since the lowest confidence levels of SCB containing the horizontal line $\hat{r} = n_1^{-1}\sum_{i=1}^{n_1}\hat{e}_{1,i}^2/n_2^{-1}\sum_{i=1}^{n_2}\hat{e}_{2,i}^2$ are 73.3%, 60.8%, 59.5% and 97.7%, respectively, where $\hat{r}$ is a consistent estimate of r under H_0, one retains the null hypothesis with the p-values = 0.267, 0.392, 0.405 and 0.023, respectively. Thus, the strata pressure variance functions over the distance interval [295.5 m, 705.1 m] of the first three pairs differ only by constant multiples while the fourth pair

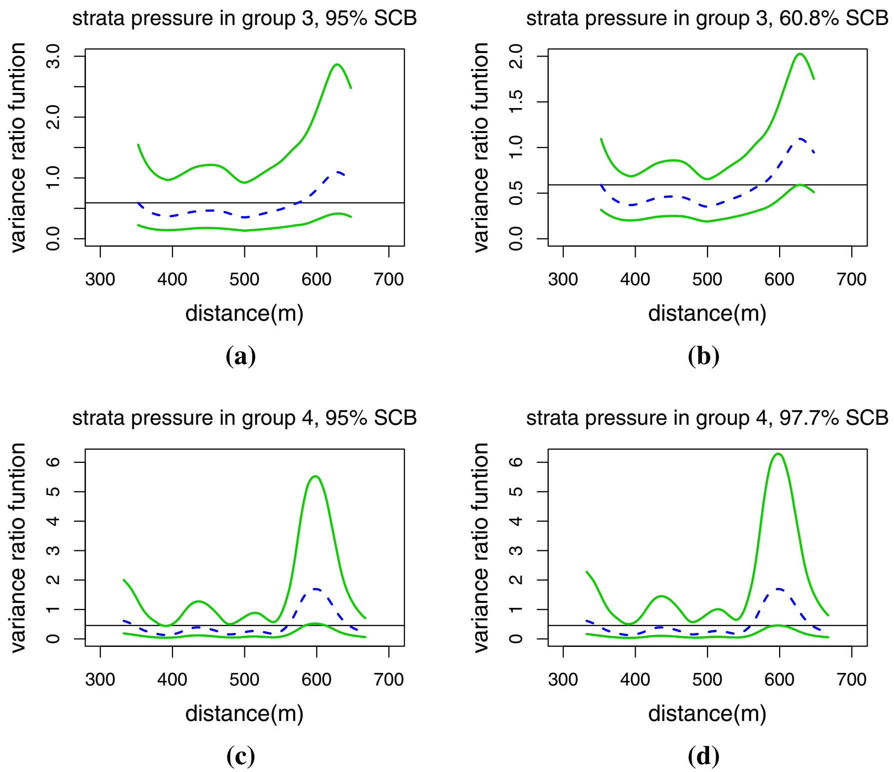

Fig. 3 Plots of the null hypothesis curve of $\hat{r} = n_1^{-1}\sum_{i=1}^{n_1}\hat{e}_{1,i}^2/n_2^{-1}\sum_{i=1}^{n_2}\hat{e}_{2,i}^2$ (solid), SCB (thick) for $\sigma_1^2(x)/\sigma_2^2(x)$ and the spline-kernel estimator $\hat{\sigma}_1^2(x)/\hat{\sigma}_2^2(x)$ (dashed), with **a** 95% SCB for pair 3; **b** lowest simultaneous confidence band containing null hypothesis for pair 3; **c** 95% SCB for pair 4; **d** lowest simultaneous confidence band containing null hypothesis for pair 4

differs significantly by a nonconstant multiple, most noticeably spiking at the location around 600 m.

According to Jiang et al. (2020), SCB for the difference $m_1(\cdot) - m_2(\cdot)$ of two mean functions are constructed only when the variance functions are proportional, i.e., when the aforementioned H_0 is not rejected. The above findings allow one to compare the mean strata pressure functions of two sites in the first three pairs by the SCB of Jiang et al. (2020), but not the two sites in the fourth pair. Such comparison provides useful safety information on relative levels of strata pressure function at various sites.

6 Conclusions

A spline-kernel estimator is proposed for the ratio of variance functions in nonparametric regression model, which is shown to be oracle efficient, that is, it uniformly approximates an infeasible estimator at the rate of $o_p\left(n_1^{-1/2} + n_2^{-1/2}\right)$. A data-driven procedure implements an asymptotical oracle SCB, which is location and scale invariant, centered around the spline-kernel variance ratio estimator, with limiting converge probability equal to that of infeasible SCB. As illustrated by strata pressure data from the Bullianta Coal Mine in Erdos City, Inner Mongolia, China, the theoretically justified SCB is a useful tool to check the ratio of variance functions in nonparametric regression, and is expected to find wide application in many scientific disciplines.

Acknowledgements This research is part of the first author's dissertation under the supervision of the second author, and has been supported exclusively by National Natural Science Foundation of China award 11771240. The authors are grateful to Professor Yaodong Jiang for providing the strata pressure data, and two Reviewers for thoughtful comments.

Appendix

The following is a reformulation of Theorems 11.1.5 and 12.3.5 in Leadbetter et al. (1983).

Lemma 1 *If a Gaussian process $\varsigma(s), 0 \le s \le T$ is stationary with mean zero and variance one, and covariance function statisfying*

$$r(t) = E\varsigma(s)\varsigma(t+s) = 1 - C|t|^{\alpha} + o\left(|t|^{\alpha}\right), \text{ as } t \to 0$$

for some constant $C > 0, 0 < \alpha \le 2$. Then as $T \to \infty$,

$$P\left[a_T\left\{\sup_{t\in[0,T]}|\varsigma(t)| - b_T\right\} \le z\right] \to e^{-2e^{-z}}, \forall z \in \mathbb{R},$$

where $a_T = (2\log T)^{1/2}$ and

$$b_T = a_T + a_T^{-1} \times \left\{\left(\frac{1}{\alpha} - \frac{1}{2}\right)\log\left(a_T^2/2\right) + \log(2\pi)^{-1/2}\left(C^{\frac{1}{\alpha}}H_{\alpha}2^{\frac{2-\alpha}{2\alpha}}\right)\right\}$$

with $H_1 = 1$, $H_2 = \pi^{-1/2}$.

Lemmas 2–4 are from Cai et al. (2019).

Lemma 2 *Under Assumption (A6), for $s = 1, 2$, as $n \to \infty$,*

$$\sup_{x\in\mathcal{I}_n}\left|\hat{f}_s(x) - 1\right| = \mathcal{O}\left(n_s^{-1}h^{-2}\right).$$

Lemma 3 *Under Assumptions (A2), (A6) and (A7), for $s = 1, 2$, as $n \to \infty$,*

$$\sup_{x\in\mathcal{I}_n}\left|A_{s,n_s}(x)\right| = \mathcal{O}\left(h^{\theta_0+p_0-1} + n_s^{-1}h^{-1}\right).$$

Lemma 4 *Under Assumptions (A2)–(A4), (A6), (A7), for $s=1,2$, as $n\to\infty$,*

$$(a)\ \sup_{x\in[0,1]}\left|B_{s,n_s}(x)-B_{s,n_s,1}(x)\right|=\mathcal{O}_p\left(n_s^{\beta_s-1}h^{-1}\right),$$
$$(b)\ \sup_{x\in[0,1]}\left|B_{s,n_s,1}(x)-B_{s,n_s,2}(x)\right|=\mathcal{O}_p\left(n_s^{-1/2}h^{1/2}\log^{1/2}n_s\right),$$
$$(c)\ \sup_{x\in\mathcal{I}_n}\left|B_{s,n_s,2}(x)-B_{s,n_s,3}(x)\right|=\mathcal{O}_p\left(n_s^{-3/2}h^{-2}\log^{1/2}n_s\right),$$
$$(d)\ \sup_{x\in[0,1]}\left|B_{s,n_s,3}(x)\right|=\mathcal{O}_p\left(n_s^{-1/2}h^{-1/2}\log^{1/2}n_s\right).$$

Denote

$$B_{n_1,n_2}(x)=\sigma_1^{-2}(x)\,B_{1,n_1}(x)-\sigma_2^{-2}(x)\,B_{2,n_2}(x)$$
$$B_{n_1,n_2,3}(x)=\sigma_1^{-2}(x)\,B_{1,n_1,3}(x)-\sigma_2^{-2}(x)\,B_{2,n_2,3}(x).$$

Lemma 5 *Under Assumptions (A2)–(A4), (A6), (A7), as $n\to\infty$,*

$$\begin{aligned}&\sup_{x\in\mathcal{I}_n}\left|\ln\frac{\tilde\sigma_1^2(x)}{\tilde\sigma_2^2(x)}-\ln\frac{\sigma_1^2(x)}{\sigma_2^2(x)}-B_{n_1,n_2,3}(x)\right|\\&\quad=\mathcal{O}_p\left(n_1^{-1/2}h^{-1/2}\log^{1/2}n_1+n_2^{-1/2}h^{-1/2}\log^{1/2}n_2\right)\\&\qquad+\mathcal{O}_p\left(h^{\theta_0+p_0-1}+n_1^{\beta_1-1}h^{-1}+n_2^{\beta_2-1}h^{-1}\right)+o_p(1).\end{aligned}$$

Consequently,

$$a_h\left\{v_n^{-1}\sup_{x\in\mathcal{I}_n}\left|\ln\frac{\tilde\sigma_1^2(x)}{\tilde\sigma_2^2(x)}-\ln\frac{\sigma_1^2(x)}{\sigma_2^2(x)}\right|\right\}=a_h\left\{v_n^{-1}\sup_{x\in\mathcal{I}_n}\left|B_{n_1,n_2,3}(x)\right|\right\}+o_p(1),$$

where a_h and v_n are given in (5).

Proof According to Lemmas 2–4, one has

$$\begin{aligned}&\sup_{x\in\mathcal{I}_n}\left|\hat f_s^{-1}(x)\left\{A_{s,n_s}(x)+B_{s,n_s}(x)\right\}\right|\\&\quad\le\sup_{x\in\mathcal{I}_n}\left|\hat f_s^{-1}(x)\right|\sup_{x\in\mathcal{I}_n}\left|A_{s,n_s}(x)+B_{s,n_s}(x)\right|\\&\quad\le\left\{1+\mathcal{O}\left(n_s^{-1}h^{-2}\right)\right\}\left\{\sup_{x\in\mathcal{I}_n}\left|A_{s,n_s}(x)\right|+\sup_{x\in\mathcal{I}_n}\left|B_{s,n_s}(x)\right|\right\}\\&\quad\le\mathcal{O}\left(h^{\theta_0+p_0-1}+n_s^{-1}h^{-1}\right)+\sup_{x\in\mathcal{I}_n}\left|B_{s,n_s}(x)\right|\\&\quad\le\mathcal{O}\left(h^{\theta_0+p_0-1}+n_s^{-1}h^{-1}\right)+\sup_{x\in[0,1]}\left|B_{s,n_s}(x)-B_{s,n_s,1}(x)\right|\end{aligned}$$

$$
\begin{aligned}
&+ \sup_{x\in[0,1]} \left|B_{s,n_s,1}(x) - B_{s,n_s,2}(x)\right| \\
&+ \sup_{x\in\mathcal{I}_n} \left|B_{s,n_s,2}(x) - B_{s,n_s,3}(x)\right| + \sup_{x\in[0,1]} \left|B_{s,n_s,3}(x)\right| \\
\le\ & \mathcal{O}_p\left(h^{\theta_0+p_0-1} + n_s^{\beta_s-1}h^{-1} + n_s^{-1/2}h^{-1/2}\log^{1/2} n_s\right)
\end{aligned}
$$

Now applying Taylor series expansions to $\ln\tilde{\sigma}_s^2(x) - \ln\sigma_s^2(x)$, for $s=1,2$

$$
\begin{aligned}
&\sup_{x\in\mathcal{I}_n} \left|\ln\tilde{\sigma}_s^2(x) - \ln\sigma_s^2(x)\right| \\
&= \sup_{x\in\mathcal{I}_n} \left|\ln\left[\sigma_s^2(x) + \hat{f}_s^{-1}(x)\left\{A_{s,n_s}(x) + B_{s,n_s}(x)\right\}\right] - \ln\sigma_s^2(x)\right| \\
&\le \sup_{x\in\mathcal{I}_n} \left|\sigma_s^{-2}(x)\hat{f}_s^{-1}(x)\left\{A_{s,n_s}(x) + B_{s,n_s}(x)\right\}\right| + o_p(1).
\end{aligned}
$$

Then one obtains

$$
\begin{aligned}
&\ln\frac{\tilde{\sigma}_1^2(x)}{\tilde{\sigma}_2^2(x)} - \ln\frac{\sigma_1^2(x)}{\sigma_2^2(x)} - B_{n_1,n_2,,3}(x) \\
&= \ln\tilde{\sigma}_1^2(x) - \ln\sigma_1^2(x) - \left\{\ln\tilde{\sigma}_2^2(x) - \ln\sigma_2^2(x)\right\} - B_{n_1,n_2,,3}(x) \\
&= \sigma_1^{-2}(x)\hat{f}_1^{-1}(x)\left\{A_{1,n_1}(x) + B_{1,n_1}(x)\right\} \\
&\quad -\sigma_2^{-2}(x)\hat{f}_2^{-1}(x)\left\{A_{2,n_2}(x) + B_{2,n_2}(x)\right\} - B_{n_1,n_2,,3}(x) + u_p(1) \\
&= \sigma_1^{-2}(x)\hat{f}_1^{-1}(x)A_{1,n_1}(x) - \sigma_2^{-2}(x)\hat{f}_2^{-1}(x)A_{2,n_2}(x) \\
&\quad +\sigma_1^{-2}(x)\left\{\hat{f}_1^{-1}(x) - 1\right\}B_{1,n_1}(x) - \sigma_2^{-2}(x)\left\{\hat{f}_2^{-1}(x) - 1\right\}B_{2,n_2}(x) \\
&\quad +B_{n_1,n_2}(x) - B_{n_1,n_2,,3}(x) + u_p(1).
\end{aligned}
$$

Since one has

$$
\begin{aligned}
&B_{n_1,n_2}(x) - B_{n_1,n_2,,3}(x) = \sigma_1^{-2}(x)B_{1,n_1}(x) - \sigma_2^{-2}(x)B_{2,n_2}(x) \\
&\ = \sigma_1^{-2}(x)\left\{B_{1,n_1}(x) - B_{1,n_1,1}(x)\right\} + \sigma_1^{-2}(x)\left\{B_{1,n_1,1}(x) - B_{1,n_1,2}(x)\right\} \\
&\quad +\sigma_1^{-2}(x)\left\{B_{1,n_1,2}(x) - B_{1,n_1,3}(x)\right\} - \sigma_2^{-2}(x)\left\{B_{2,n_2}(x) - B_{2,n_2,1}(x)\right\} \\
&\quad -\sigma_2^{-2}(x)\left\{B_{2,n_2,1}(x) - B_{2,n_2,2}(x)\right\} - \sigma_2^{-2}(x)\left\{B_{2,n_2,2}(x) - B_{2,n_2,3}(x)\right\},
\end{aligned}
\tag{14}
$$

and according to Lemmas 2–4, one has

$$
\sup_{x\in\mathcal{I}_n}\left|\sigma_s^{-2}(x)\hat{f}_s^{-1}(x)A_{s,n_s}(x)\right| = \mathcal{O}\left(h^{\theta_0+p_0-1} + n_s^{-1}h^{-1}\right), \tag{15}
$$

$$\sup_{x\in\mathcal{I}_n}\left|\sigma_s^{-2}(x)\left\{\hat{f}_s^{-1}(x)-1\right\}B_{1,n_s}(x)\right|=\mathcal{O}_p\left(n_s^{\beta_s-1}h^{-1}+n_s^{-1/2}h^{-1/2}\log^{1/2}n_s\right), \tag{16}$$

Hence combining (14), (15) and (16), the proof is completed.

Denote the following processes

$$\begin{aligned}
Y_{1,n_1,1}(x)&=h^{-1}n_1^{-1/2}\left(\mu_{1,4}-1\right)^{1/2}\int K(x-u/h)\,dW_{1,n_1}(u),\ x\in\left[1,h^{-1}-1\right],\\
Y_{2,n_2,1}(x)&=h^{-1}n_2^{-1/2}\left(\mu_{2,4}-1\right)^{1/2}\int K(x-u/h)\,dW_{2,n_2}(u),\ x\in\left[1,h^{-1}-1\right],\\
Y_{1,n_1,2}(x)&=h^{-1/2}n_1^{-1/2}\left(\mu_{1,4}-1\right)^{1/2}\int K(x-r)\,dW_{1,n_1}(r),\ x\in\left[1,h^{-1}-1\right],\\
Y_{2,n_2,2}(x)&=h^{-1/2}n_2^{-1/2}\left(\mu_{2,4}-1\right)^{1/2}\int K(x-r)\,dW_{2,n_2}(r),\ x\in\left[1,h^{-1}-1\right].
\end{aligned}$$

As $\mathrm{E}\left\{B_{n_1,n_2,3}^2(x)\right\}=h^{-1}\left\{n_1^{-1}\left(\mu_{1,4}-1\right)+n_2^{-1}\left(\mu_{2,4}-1\right)\right\}\int_{-1}^1K^2(u)\,du$, one obtains the following standard Gaussian processes,

$$\triangle_1(x)=\frac{B_{n_1,n_2,3}(x)}{h^{-1/2}\left[\left\{n_1^{-1}\nu_{1,4}+n_2^{-1}\nu_{2,4}\right\}\int_{-1}^1K^2(u)\,du\right]^{1/2}},\ x\in[h,1-h], \tag{17}$$

$$\triangle_2(x)=\frac{Y_{1,n_1,1}(x)-Y_{2,n_2,1}(x)}{h^{-1/2}\left[\left\{n_1^{-1}\nu_{1,4}+n_2^{-1}\nu_{2,4}\right\}\int_{-1}^1K^2(u)\,du\right]^{1/2}},\ x\in\left[1,h^{-1}-1\right], \tag{18}$$

where $\nu_{1,4}=\mu_{1,4}-1$ and $\nu_{2,4}=\mu_{2,4}-1$.

Another standard Gaussian process is

$$\frac{Y_{1,n_1,2}(x)-Y_{2,n_2,2}(x)}{h^{-1/2}\left[\left\{n_1^{-1}\nu_{1,4}+n_2^{-1}\nu_{2,4}\right\}\int_{-1}^1K^2(u)\,du\right]^{1/2}},\ x\in\left[1,h^{-1}-1\right],$$

which is $\zeta(x)$ defined in (10).

Lemma 6 *The absolute maximum of the process $\triangle_1(x)$ follows the same as that of $\triangle_2(x)$, and the absolute maximum of the process $\triangle_2(x)$ follows the same as that of $\zeta(x)$, that is*

$$\sup_{x\in[h,1-h]}\left|\triangle_1(x)\right|\stackrel{d}{=}\sup_{x\in\left[1,h^{-1}-1\right]}\left|\triangle_2(x)\right|\stackrel{d}{=}\sup_{x\in\left[1,h^{-1}-1\right]}\left|\zeta(x)\right|.$$

Proof This lemma can be easily obtained by noting the fact that for $s=1,2$, the process $B_{n_1,n_2,3}(x), x\in[h,1-h]$ has the same probability law as $Y_{1,n_1,1}(x)-Y_{2,n_2,1}(x), x\in\left[1,h^{-1}-1\right]$, and the process $Y_{s,n_s,1}(x), x\in[h,1-h]$ has the same probability law as $Y_{s,n_s,2}(x), x\in\left[1,h^{-1}-1\right]$.

Proof of Proposition 1 Proposition 1 is a direct corollary of Lemma 5, Lemma 6 and Proposition 3.

Proof of Proposition 2 According to Theorem 2 in Cai et al. (2019) and applying Taylor expansion, one has

$$
\begin{aligned}
&\sup_{x\in\mathcal{I}_n}\left|\ln\frac{\hat{\sigma}_1^2(x)}{\hat{\sigma}_2^2(x)}-\ln\frac{\tilde{\sigma}_1^2(x)}{\tilde{\sigma}_2^2(x)}\right| = \sup_{x\in\mathcal{I}_n}\left|\ln\hat{\sigma}_1^2(x)-\ln\tilde{\sigma}_1^2(x)-\left\{\ln\hat{\sigma}_2^2(x)-\ln\tilde{\sigma}_2^2(x)\right\}\right| \\
&\le \sup_{x\in\mathcal{I}_n}\left|\ln\hat{\sigma}_1^2(x)-\ln\tilde{\sigma}_1^2(x)\right| + \sup_{x\in\mathcal{I}_n}\left|\ln\hat{\sigma}_2^2(x)-\ln\tilde{\sigma}_2^2(x)\right| \\
&= \sup_{x\in\mathcal{I}_n}\left|\tilde{\sigma}_1^{-2}(x)\left\{\hat{\sigma}_1^2(x)-\tilde{\sigma}_1^2(x)\right\}\right| + \sup_{x\in\mathcal{I}_n}\left|\tilde{\sigma}_2^{-2}(x)\left\{\hat{\sigma}_2^2(x)-\tilde{\sigma}_2^2(x)\right\}\right| + \mathcal{O}_p(n_1^{-1}+n_2^{-1}) \\
&\le c_\sigma^{-2}\sup_{x\in\mathcal{I}_n}\left|\hat{\sigma}_1^2(x)-\tilde{\sigma}_1^2(x)\right| + c_\sigma^{-2}\sup_{x\in\mathcal{I}_n}\left|\hat{\sigma}_2^2(x)-\tilde{\sigma}_2^2(x)\right| + \mathcal{O}_p(n_1^{-1}+n_2^{-1}) = o_p(n^{-1/2}),
\end{aligned}
$$

which completes the proof.

Proof of Proposition 3 For Gaussian process $\zeta(x)$, its correlation function is

$$
\begin{aligned}
r(x-y) = \operatorname{corr}(\zeta(x),\zeta(y)) &= \frac{\mathrm{E}\{\zeta(x)\zeta(y)\}}{\operatorname{var}^{1/2}\{\zeta(x)\}\operatorname{var}^{1/2}\{\zeta(y)\}} \\
&= \frac{\left(n_1^{-1}\nu_{1,4}+n_2^{-1}\nu_{2,4}\right)(K*K)(x-y)}{\left(n_1^{-1}\nu_{1,4}+n_2^{-1}\nu_{2,4}\right)\int_{-1}^1 K^2(u)\,du} \\
&= \frac{(K*K)(x-y)}{\int_{-1}^1 K^2(u)\,du},
\end{aligned}
$$

which implies that

$$
r(t) = \frac{\int K(u)K(u-t)\,du}{\int_{-1}^1 K^2(u)\,du}.
$$

Define next a Gaussian process $\varsigma(t), 0 \le t \le T = T_n = h^{-1}-2$,

$$
\varsigma(t) = \zeta(t+1)\left\{\int_{-1}^1 K^2(u)\,du\right\}^{-1/2},
$$

which is stationary with mean zero and variance one, and covariance function

$$
r(t) = \mathrm{E}\varsigma(s)\varsigma(t+s) = 1 - Ct^2 + o\left(t^2\right), t \to 0,
$$

with $C = \int_{-1}^{1} K^{(1)}(u)^2\, du / 2 \int_{-1}^{1} K^2(u)\, du$. Hence applying Lemmas 1–6, one has for $h \to 0$ or $T \to \infty$,

$$\mathrm{P}\left[a_T \left\{\sup_{t\in[0,T]} |\varsigma(t)| - b_T\right\} \le z\right] \to e^{-2e^{-z}}, \forall z \in \mathbb{R},$$

where $a_T = (2\log T)^{1/2}$ and $b_T = a_T + a_T^{-1}\left\{2^{-1}\log\left(C_K/\left(4\pi^2\right)\right)\right\}$. Note that

$$a_h a_T^{-1} \to 1, a_T(b_T - b_h) \to 0.$$

Hence, applying Slutsky's Theorem twice, one obtains that

$$a_h\left\{\sup_{t\in[0,T]} |\varsigma(t)| - b_h\right\} = a_h a_T^{-1}\left[a_T\left\{\sup_{t\in[0,T]} |\varsigma(t)| - b_T\right\}\right] + a_h(b_T - b_h),$$

which converges in distribution to the same limit as $a_T\left\{\sup_{t\in[0,T]} |\varsigma(t)| - b_T\right\}$. Thus

$$\mathrm{P}\left[a_h\left\{\sup_{s\in[1,h^{-1}-1]} |\zeta(s)| - b_h\right\} < z\right] \to \exp\{-2\exp(-z)\}, z \in \mathbb{R}.$$

Hence the proof is completed.

Proof of Theorem 1 According to Proposition 1, as $n \to \infty$,

$$\mathbb{P}\left[a_h\left\{v_n^{-1}\sup_{x\in\mathcal{I}_n}\left|\ln\frac{\tilde{\sigma}_1^2(x)}{\tilde{\sigma}_2^2(x)} - \ln\frac{\sigma_1^2(x)}{\sigma_2^2(x)}\right| - b_h\right\} \le z\right] \to \exp\{-2\exp(-z)\}, z \in \mathbb{R}, \tag{19}$$

where a_h, b_h and v_n are given in (5). Finally applying Proposition 2, one obtains

$$a_h\left\{v_n^{-1}\sup_{x\in\mathcal{I}_n}\left|\ln\frac{\hat{\sigma}_1^2(x)}{\hat{\sigma}_2^2(x)} - \ln\frac{\tilde{\sigma}_1^2(x)}{\tilde{\sigma}_2^2(x)}\right|\right\} = o_p\left(\left\{\log\left(h^{-1}\right)\right\}^{1/2} h^{1/2}\right) = o_p(1).$$

Using Slutsky's Theorem one can substitute $\ln\frac{\hat{\sigma}_1^2(x)}{\hat{\sigma}_2^2(x)}$ for $\ln\frac{\tilde{\sigma}_1^2(x)}{\tilde{\sigma}_2^2(x)}$ in (19). Hence the proof of Theorem 1 is completed.

References

Bose S, Mahalanobis P (1935) On the distribution of the ratio of variances of two samples drawn from a given normal bivariate correlated population. Sankhyā Indian J Stat 2:65–72

Bickel P, Rosenblatt M (1973) On some global measures of deviations of density function estimates. Ann Stat 31:1852–1884

Brown L, Levine M (2007) Variance estimation in nonparametric regression via the difference sequence method. Ann Stat 35:2219–2232

Claeskens G, Van Keilegom I (2003) Bootstrap confidence bands for regression curves and their derivatives. Ann Stat 31:1852–1884

Cai L, Yang L (2015) A smooth simultaneous confidence band for conditional variance function. TEST 24:632–655

Cai L, Liu R, Wang S, Yang L (2019) Simutaneous confidence bands for mean and varience functions based on deterministic design. Stat Sin 29:505–525

Cai L, Li L, Huang S, Ma L, Yang L (2020) Oracally efficient estimation for dense functional data with holiday effects. TEST 29:282–306

Cao G, Wang L, Li Y, Yang L (2016) Oracle-efficient confidence envelopes for covariance functions in dense functional data. Stat Sin 26:359–383

Cao G, Yang L, Todem D (2012) Simultaneous inference for the mean function based on dense functional data. J Nonparametr Stat 24:359–377

Degras D (2011) Simultaneous confidence bands for nonparametric regression with functional data. Stat Sin 21:1735–1765

de Boor C (2001) A Practical Guide to Splines. Springer, New York

Eubank R, Speckman P (1993) Confidence bands in nonparametric regression. J Am Stat Assoc 88:1287–1301

Finney D (1938) The distribution of the ratio of estimates of the two variances in a sample from a normal bi-variate population. Biometrika 30:190–192

Fisher R (1924) On a distribution yielding the error function of several well known statistics. Proc Int Congr Math 2:805–813

Fan J, Gijbels I (1996) Local Polynomial Modelling and Its Applications. Chapman and Hall, London

Gayen A (1950) The distribution of the variance ratio in random samples of any size drawn from non-normal universes. Biometrika 37:236–255

Gu L, Wang L, Härdle W, Yang L (2014) A simultaneous confidence corridor for varying coefficient regression with sparse functional data. TEST 23:806–843

Gu L, Yang L (2015) Oracally efficient estimation for single-index link function with simultaneous confidence band. Electron J Stat 9:1540–1561

Gu L, Wang S, Yang L (2019) Simultaneous confidence bands for the distribution function of a finite population in stratified sampling. Ann Inst Stat Math 71:983–1005

Härdle W (1989) Asmptotic maximal deviation of M-smoothers. J Multivar Anal 29:163–179

Härdle W, Marron J (1991) Bootstrap simultaneous error bars for nonparametric regression. Ann Stat 19:778–796

Hall P, Titterington D (1988) On confidence bands in nonparametric density estimation and regression. J Multivar Anal 27:228–254

James G (1951) The comparison of several groups of observations when the ratios of the population variances are unknown. Biometrika 38:324–329

Jiang J, Cai L, Yang L (2020) Simultaneous confidence band for the difference of regression functions of two samples. Commun Stat Theory Methods. https://doi.org/10.1080/03610926.2020.1800039

Johnston G (1982) Probabilities of maximal deviations for nonparametric regression function estimates. J Multivar Anal 12:402–414

Ju J, Xu J (2013) Structural characteristics of key strata and strata behaviour of a fully mechanized longwall face with 7.0m height chocks. Int J Rock Mech Min Sci 58:46–54

Leadbetter MR, Lindgren G, Rootzén H (1983) Extremes and related properties of random sequences and processes. Springer, New York

Levine M (2006) Bandwidth selection for a class of difference-based variance estimators in the nonparametric regression: a possible approach. Comput Stat Data Anal 50:3405–3431

Ma S, Yang L, Carroll R (2012) A simultaneous confidence band for sparse longitudinal regression. Stat Sin 22:95–122

Qian M, Shi P, Xu J (2010) Mining pressure and strata control. China University of Mining and Technology Press, Beijing

Scheffé H (1942) On the ratio of the variances of two normal populations. Ann Math Stat 13:371–388

Song Q, Yang L (2009) Spline confidence bands for variance functions. J Nonparametr Stat 5:589–609

Song Q, Liu R, Shao Q, Yang L (2014) A simultaneous confidence band for dense longitudinal regression. Commun Stat Theory Methods 43:5195–5210

Wang J, Yang L (2009) Polynomial spline confidence bands for regression curves. Stat Sin 19:325–342
Wang J (2012) Modelling time trend via spline confidence band. Ann Inst Stat Math 64:275–301
Wang J, Liu R, Cheng F, Yang L (2014) Oracally efficient estimation of autoregressive error distribution with simultaneous confidence band. Ann Stat 42:654–668
Wang J, Wang S, Yang L (2016) Simultaneous confidence bands for the distribution function of a finite population and its superpopulation. TEST 25:692–709
Welch B (1938) The significance of the difference between two means when the population variances are unequal. Biometrika 29:350–362
Xia Y (1998) Bias-corrected confidence bands in nonparametric regression. J R Stat Soc Ser B 60:797–811
Xue L, Yang L (2006) Additive coefficient modeling via polynomial spline. Stat Sin 16:1423–1446
Zhao Z, Wu W (2008) Confidence bands in nonparametric time series regression. Ann Stat 36:1854–1878
Zhang Y, Yang L (2018) A smooth simultaneous confidence band for correlation curve. TEST 27:247–269
Zheng S, Yang L, Härdle W (2014) A smooth simultaneous confidence corridor for the mean of sparse functional data. J Am Stat Assoc 109:661–673
Zheng S, Liu R, Yang L, Härdle W (2016) Statistical inference for generalized additive models:simultaneous confidence corridors and variable selection. TEST 25:607–626