Ann. Inst. Statist. Math.
Vol. 46, No. 2, 251–265 (1994)

TESTING FOR NO EFFECT IN NONPARAMETRIC REGRESSION VIA SPLINE SMOOTHING TECHNIQUES

JUEI-CHAO CHEN

Graduate Institute of Statistics and Actuarial Science, Feng Chia University, Taichung 40724, Taiwan, ROC

(Received June 8, 1992; revised August 2, 1993)

Abstract. We propose three statistics for testing that a predictor variable has no effect on the response variable in regression analysis. The test statistics are integrals of squared derivatives of various orders of a periodic smoothing spline fit to the data. The large sample properties of the test statistics are investigated under the null hypothesis and sequences of local alternatives and a Monte Carlo study is conducted to assess finite sample power properties.

Key words and phrases: Asymptotic distribution, local alternatives, nonparametric regression, Monte Carlo.

1. Introduction

Regression analysis is used to study relationships between the response variable and a predictor variable; therefore, all of the inference is based on the assumption that the response variable actually depends on the predictor variable. Testing for no effect is the same as checking this assumption. In this paper, we propose three statistics for testing that a predictor variable has no effect on the response variable and derive their asymptotic distributions.

Suppose we have an experiment which yields observations $(t_{1n}, y_1), \ldots, (t_{nn}, y_n)$, where y_j represents the value of a response variable y at equally spaced values $t_{jn} = (j-1)/n$, $j = 1, \ldots, n$, of the predictor variable t. The regression model is

$$y_j = \mu(t_{jn}) + \epsilon_j, \qquad j = 1, \ldots, n, \tag{1.1}$$

where the ϵ_j are iid random errors with $E[\epsilon_1] = 0$, $\mathrm{Var}[\epsilon_1] = \sigma^2$ and $0 < E[\epsilon_1^4] < \infty$. We assume that $\mu \in W_{2,\mathrm{per}}^2[0,1] = \{g : g^{(j)}$ is absolutely continuous, $j = 0, 1$, $\int_0^1 g^{(2)}(t)^2 dt < \infty$ and $g^{(j)}(0) = g^{(j)}(1)$, $j = 0, 1\}$. The periodicity conditions imposed on μ allow us to use Fourier analysis techniques that simplify subsequent comparisons and mathematical developments. For notational convenience, we will also take n to be odd throughout the paper.

Testing for no effect means testing that the regression function is a constant. Hence, we rewrite model (1.1) as

$$y_j = \beta + f(t_{jn}) + \epsilon_j, \tag{1.2}$$

$j = 1, \ldots, n$, where β is an unknown constant and f in (1.2) is an unknown function that can be assumed, without loss of generality, to satisfy $\int_0^1 f(t)dt = 0$. Therefore, the null hypothesis of no effect is equivalent to $H_0 : f = 0$. Since $\mu \in W^2_{2,\text{per}}[0,1]$, and f is orthogonal to the unit function, $f = 0$ if and only if, $\int_0^1 f^{(m)}(t)^2 dt = 0$, $m = 0$, 1 or 2. Consequently, one could test H_0 using statistics of the form $\int_0^1 \hat{f}^{(m)}(t)^2 dt$, $m = 0$, 1 and 2 where $\hat{f}$ is some estimator of f. This is what we propose to do.

As an estimator of f, we will use a periodic smoothing spline. Specifically, let $r_j = y_j - \bar{y}$. Then, a smoothing spline estimator of f is the minimizer of

$$\frac{1}{n}\sum_{j=1}^{n}(r_j - g(t_{jn}))^2 + \lambda \int_0^1 [g^{(2)}(t)]^2 dt, \qquad \lambda > 0, \tag{1.3}$$

over all $g \in W^2_{2,\text{per}}[0,1]$. The term $(1/n)\sum_{j=1}^{n}(r_j - g(t_{jn}))^2$ in (1.3) measures goodness of fit, while $\int_0^1 g^{(2)}(t)^2 dt$ measures smoothness. The smoothing parameter λ in (1.3) controls the tradeoff between variance and bias of the estimator.

The minimizer of (1.3) is well approximated by (cf. Rice and Rosenblatt (1981) or Eubank (1988), p. 304)

$$f_{\lambda p}(t) = \sum_{|j| \le (n-1)/2}{}' \frac{\tilde{a}_{jn} e^{2\pi i j t}}{1 + \lambda (2\pi j)^4}, \tag{1.4}$$

where the $\sum'$ indicates summation excluding the zero index and the $\tilde{a}_{jn}$ are the discrete sample Fourier coefficients defined as

$$\tilde{a}_{jn} = n^{-1} \sum_{k=1}^{n} y_k e^{-2\pi i j((k-1)/n)}, \qquad |j| \le \frac{n-1}{2}. \tag{1.5}$$

We use (1.4) as our estimator of f in what follows and refer to it as a periodic smoothing spline.

There are a number of tests available for the no effect hypothesis or, more generally, for testing goodness of fit of a linear model. For example, Graybill (1976) or Kleinbaum *et al.* (1988) discuss the classical parametric methodology for this purpose. Recently, Cox *et al.* (1988), Cox and Koh (1989) and Buckley (1991) applied a Bayesian approach to construct test statistics. Tests derived by von Neumann (1941), Härdle and Mammen (1988), Munson and Jernigan (1989), Eubank and Spiegelman (1990), Jayasuriya (1990), Raz (1990), Staniswalis and Severini (1991), Müller (1992), Eubank and Hart (1993), and Eubank and LaRiccia (1993) use various nonparametric smoothing techniques to find an estimator of f and construct a test statistic for H_0. Procedures which employ the smoothing parameter as a test statistic have been studied by Eubank and Hart (1992) and Hart and Wehrly (1992).

The test statistics we will consider are

$$T_{mp} = \int_0^1 f_{\lambda p}^{(m)}(t)^2 dt, \qquad m = 0,\ 1 \text{ and } 2.$$

When $m = 0$ this statistic has been studied by Eubank and LaRiccia (1993) under the assumption of normal errors. For $m = 1$ and 2, the T_{mp} provide new proposals. In the next section, we study the asymptotic distribution theory of these test statistics under both the null hypothesis and local alternatives. The results of a small simulation to study the power properties of the test statistics are reported in Section 3. Proofs of the theorems are then collected in the Appendix.

2. Asymptotic distribution theory

In this section, we study the large sample properties of the T_{mp} for $m = 0$, 1 and 2. Initially, we assume here that σ^2 is known and discuss how this assumption can be relaxed subsequently. We begin with some notational preliminaries.

Let $*$ denote complex conjugation and for any functions g and h belonging to $L_2[0,1] = \{f : \int_0^1 |f(t)|^2 dt < \infty\}$ with $|f(t)|^2 = f(t)f^*(t)$, define the inner product $\langle g, h\rangle = \int_0^1 g(t)h^*(t)dt$ and norm $\|g\| = \langle g, g\rangle^{1/2}$. Setting $x_j(t) = e^{2\pi ijt}$, the j-th Fourier coefficient of a function f is then defined to be $a_j = \langle f, x_j\rangle = \int_0^1 f(t)e^{-2\pi ijt}dt$.

For $m = 0$, $T_{0p} = \int_0^1 [f_{\lambda p}(t)]^2 dt = \sum'_{|j|\le(n-1)/2} |\tilde{a}_{jn}|^2/(1+\lambda(2\pi j)^4)^2$. The specific test statistic we consider is a standardization of T_{0p}; namely,

$$Z_{0p} = \frac{T_{0p} - \dfrac{\sigma^2}{n}\sum'_{|j|\le(n-1)/2}\dfrac{1}{(1+\lambda(2\pi j)^4)^2}}{\dfrac{\sigma^2}{n}\sqrt{2\sum'_{|j|\le(n-1)/2}\dfrac{1}{(1+\lambda(2\pi j)^4)^4}}}. \tag{2.1}$$

Theorem 2.1 below states that Z_{0p} has an asymptotic normal distribution.

THEOREM 2.1. *Assume that $n \to \infty$, $\lambda \to 0$ in such a way that $n\lambda^{1/2} \to \infty$. Suppose $f = h(n)g$ with $g \in W_{2,\text{per}}^2[0,1]$ and $h(n) = n^{-1/2}\lambda^{-1/16}$. Then Z_{0p} converges in distribution to a $N(\|g\|^2/\sigma^2\sqrt{2C_0}, 1)$ random variable for $C_0 = \pi^{-1}\int_0^\infty 1/(1+x^4)^4 dx$.*

Theorem 2.1 generalizes results in Eubank and LaRiccia (1993) to the case of nonnormal error distributions. It has the implication that a test of the above form can detect local alternatives converging to the null at the rate $n^{-1/2}\lambda^{-1/16}$ or slower. For any fixed alternative, it can be shown that Z_{0p} also provides a consistent test.

For $m = 1$ we obtain a new test statistic $T_{1p} = \int_0^1 [f_{\lambda p}^{(1)}(t)]^2 dt = \sum'_{|j|\le(n-1)/2} (2\pi j)^2 |\tilde{a}_{jn}|^2/(1+\lambda(2\pi j)^4)^2$ whose standardized form is

$$Z_{1p} = \frac{T_{1p} - \dfrac{\sigma^2}{n}\sum'_{|j|\le(n-1)/2} \dfrac{(2\pi j)^2}{(1+\lambda(2\pi j)^4)^2}}{\dfrac{\sigma^2}{n}\sqrt{2\sum'_{|j|\le(n-1)/2} \dfrac{(2\pi j)^4}{(1+\lambda(2\pi j)^4)^4}}}. \tag{2.2}$$

THEOREM 2.2. *Assume that $n \to \infty$, $\lambda \to 0$ in such a way that $n\lambda^{5/8} \to \infty$. Suppose $f = h(n)g$ with $g \in W_{2,\mathrm{per}}^2[0,1]$ and $h(n) = n^{-1/2}\lambda^{-5/16}$. Then, Z_{1p} converges in distribution to a $N(\|g^{(1)}\|^2/\sigma^2\sqrt{2C_1}, 1)$ random variable for $C_1 = \pi^{-1}\int_0^\infty x^4/(1+x^4)^4 dx$.*

Finally, we have the statistic for $m = 2$, $T_{2p} = \int_0^1 [f_{\lambda p}^{(2)}(t)]^2 dt = \sum'_{|j|\le(n-1)/2} (2\pi j)^4 |\tilde{a}_{jn}|^2/(1+\lambda(2\pi j)^4)^2$ which, when recentered and rescaled, becomes

$$Z_{2p} = \frac{T_{2p} - \dfrac{\sigma^2}{n}\sum'_{|j|\le(n-1)/2} \dfrac{(2\pi j)^4}{(1+\lambda(2\pi j)^4)^2}}{\dfrac{\sigma^2}{n}\sqrt{2\sum'_{|j|\le(n-1)/2} \dfrac{(2\pi j)^8}{(1+\lambda(2\pi j)^4)^4}}}. \tag{2.3}$$

THEOREM 2.3. *Assume that $n \to \infty$, $\lambda \to 0$ in such a way that $n\lambda^{9/8} \to \infty$. Suppose $f = h(n)g$ with $g \in W_{2,\mathrm{per}}^2[0,1]$ and let b_j be the j-th Fourier coefficient of g. Further assume that $|b_j|^2 \sim j^{-(5+\delta)}$ for some $\delta > 1$. Then, if $h(n) = n^{-1/2}\lambda^{-9/16}$, Z_{2p} converges in distribution to a $N(\|g^{(2)}\|^2/\sigma^2\sqrt{2C_2}, 1)$ random variable for $C_2 = \pi^{-1}\int_0^\infty x^8/(1+x^4)^4 dx$.*

Remark 1. Theorems 2.2 and 2.3 have the implication that Z_{1p} and Z_{2p} both have asymptotic standard normal distributions under the null hypothesis. Our local alternative analysis also reveals that Z_{1p} can detect alternatives converging to the null at the rate $n^{-1/2}\lambda^{-5/16}$ or slower while Z_{2p} can detect local alternatives converging as fast as $n^{-1/2}\lambda^{-9/16}$. Like the Z_{0p} based test, the ones using either Z_{1p} or Z_{2p} are consistent against any fixed alternative.

Remark 2. Although Theorems 2.1–2.3 suggest that suitable critical values for the Z_{mp}, $m = 0, 1, 2$ can be obtained from the standard normal distribution, results in Jayasuriya (1990) and Eubank and LaRiccia (1993) indicate that such normal approximations are typically not adequate even for fairly large samples. The problem is that the test statistics behave like weighted sums of chi-square random variable, in an asymptotic sense, and accordingly approach normality quite slowly. An alternative approximation developed to deal with such cases (i.e., weighted sums of chi-squares that are asymptotically normal) has been proposed by Buckley and Eagleson (1988). Using their approach we can approximation the $100(1-\alpha)$-th percentile of Z_{mp}, $m = 0, 1, 2$, by $(\chi^2_{d,1-\alpha} - d)/\sqrt{2d}$, where $d = [\sum'_{|j|\leq(n-1)/2} E^2_{mj}]^3/[\sum'_{|j|\leq(n-1)/2} E^3_{mj}]^2$, the E_{mj} are given in (A1)–(A3) of the Appendix and $\chi^2_{d,1-\alpha}$ denotes the $100(1-\alpha)$-th percentile of a chi-squared distribution with d degrees of freedom. This method produced satisfactory results in the simulations in Section 3.

Remark 3. We assumed that σ^2 in (2.1), (2.2) and (2.3) was known in establishing the asymptotic distributions for the Z_{mp}, $m = 0, 1, 2$. However, Theorems 2.1–2.3 remain valid if σ^2 is replaced by any $\sqrt{n}$-consistent estimator. Examples of such estimators can be found in Gasser *et al.* (1986) and Hall *et al.* (1990).

3. Power properties

The power properties of our testing procedures will be examined in this section. We begin with a discussion of asymptotic power.

For a given level α and alternative g, we can use Theorems 2.1–2.3 to get expressions for the large sample powers of our tests. Let $P(A \mid f)$ denote the conditional probability of an event A under the alternative $\mu = \beta + f$. Then, we have the following result.

THEOREM 3.1. *Assume the conditions of Theorems* 2.1–2.3 *hold and let* $h(n) = n^{-1/2}\lambda^{-(4m+1)/16}$. *Then, for any given* $\alpha \in (0,1)$ *and* $0 \leq \|g^{(m)}\|^2 < \infty$,

$$\lim_{n\to\infty} P(Z_{mp} \geq Z_{1-\alpha} \mid h(n)g) = 1 - \Phi\left(Z_{1-\alpha} - \frac{\|g^{(m)}\|^2}{\sigma^2\sqrt{2C_m}}\right),$$

where $Z_{1-\alpha}$ *is the* $100(1-\alpha)$*-th percentile of the standard normal distribution.*

If we assume that $f(0) = 0$, then $f(t) = \int_0^t f^{(1)}(u)du$ and the Cauchy-Schwarz inequality gives $\int_0^1 f(t)^2 dt \leq \|f^{(1)}\|^2 \int_0^1 t dt = .5\|f^{(1)}\|^2$. Similarly, if $f^{(1)}(0) = 0$ we obtain $\|f\|^2 \leq .5\|f^{(1)}\|^2 \leq .25\|f^{(2)}\|^2$. One finds that $C_0 \doteq .265$ and $C_1 = C_2 \doteq .088$. Consequently, Theorem 3.1 has the implication that the asymptotic power of T_{mp} is an increasing function of m in this case.

Cox and Koh (1989) and Buckley (1991) propose tests based on Bayesian methodology that can also be adapted for use in our setting. Basically, one fits models the regression function as $\beta + kW(\cdot)$, where W is a zero mean, normal process and then derives a locally most powerful test for the hypothesis that $k = 0$, which is equivalent to our H_0. If we choose the covariance kernel for W to be $B(s,t) = \sum'_j (1/(2\pi j)^4)e^{2\pi i(s-t)j}$, we obtain a parallel of the Cox/Koh test for the no effect hypothesis. The resulting test statistic is well approximated by

$$T_{CK} = \sum_{|j|\leq(n-1)/2}{}' \frac{|\tilde{a}_{jn}|^2}{\sigma^2(2\pi j)^4}. \tag{3.1}$$

This resembles the Cox and Koh (1989) statistic for testing goodness-of-fit of a polynomial model given in equation (2.13) of their paper. Our problem differs somewhat from theirs since we are concerned with testing for no effect with a periodic regression function. This requires some minor alterations in the Bayesian modeling. In particular, we must use a different covariance kernel for our Gaussian process which is what causes the differences in the actual form of the test statistic.

The Fourier representation (3.1) for T_{CK} makes it possible to affect some direct qualitative comparisons with the statistics in Section 2. First, observe that $\tilde{a}_{jn}$ in (1.5) provides an estimator of a_j, the j-th Fourier coefficient of f. If f is a high frequency function then $|a_j|$ will tend to be large for larger values of $|j|$. Notice that the test statistics T_{CK}, T_{0p}, T_{1p} and T_{2p} all rely on the sample Fourier coefficients, but have different weights. T_{CK} downweights $|\tilde{a}_{jn}|$ as j increases, while T_{0p} uses roughly uniform weights for $j < (2\pi\lambda^{1/4})^{-1}$ and downweights for larger j. T_{1p} and T_{2p} give larger weights to $|\tilde{a}_{jn}|^2$ for large $|j|$ than T_{CK} and T_{0p}. We, therefore, expect the test statistics T_{1p} and T_{2p} to have more power to detect high frequency alternatives than T_{CK} and T_{0p}.

The goal of the remainder of the paper is to ascertain the extent that our Fourier analysis intuition is realized in quantitative comparisons. In this regard, one may use Theorem 3.1 to prove parallels of the Corollary and Theorem 2 in Eubank and LaRiccia (1993) that provide analytic (asymptotic) comparisons of T_{CK} and the T_{mp}. We will not pursue that here but instead conduct more direct comparisons using Monte Carlo techniques.

To ascertain how well our test might work in finite samples with fixed alternatives we conducted a small scale simulation experiment. Normal errors were used to generate random samples of size 101 from model (1.2). Without loss of generality, we took $\beta = 0$. The error variance was assumed known and equal to 1 and the design points t_{jn} were chosen to be equally spaced over $[0,1]$. For the function f in (1.2) we chose $f(t) = \rho\cos(2\pi\nu t)$. Since $\sigma = 1$, the value used for ρ can be regarded as a signal to noise ratio with smaller values indicating increasing difficulty in estimation. We considered the specific choices $\rho = 0.25,\ 0.5,\ 1.0$ and 1.5 in our power study. For ν we used $\nu = 1,\ 3,\ 9$ and 12 so that the alternatives f will be of higher frequency as ν increases.

All three of the tests T_{0p}, T_{1p} and T_{2p} depend on a smoothing parameter λ. We used integrated mean squared error optimal choices for λ in the simulation. More specifically, one may show (Chen (1992)) that the value of λ which minimizes the integrated mean squared error for estimating f by $f_{\lambda p}$ is, approximately, $\lambda_{0,\text{opt}} = ((\sigma^2 C_0)/(4n\|f^{(2)}\|^2))^{4/5}$ if $f \in W^2_{2,\text{per}}[0,1]$. Under this same restriction, the asymptotically optimal choice for λ when estimating $f^{(1)}$ by $f^{(1)}_{\lambda p}$ is $\lambda_{1,\text{opt}} = ((3\sigma^2 C_1)/(2n\|f^{(2)}\|^2))^{4/5}$. Assuming $f \in W^2_{3,\text{per}}[0,1]$, the asymptotically optimal λ for estimating $f^{(2)}$ by $f^{(2)}_{\lambda p}$ is $\lambda_{2,\text{opt}} = ((5\sigma^2 C_2)/(2n\|f^{(3)}\|^2))^{4/7}$. For our choice of $f(t) = \rho\cos(2\pi\nu t)$, this translates into

$$\lambda_{0,\text{opt}} = \left[\frac{0.265165}{32n\rho^2(\nu\pi)^4}\right]^{4/5}, \qquad \lambda_{1,\text{opt}} = \left[\frac{0.265165}{16n\rho^2(\nu\pi)^4}\right]^{4/5} \qquad \text{and}$$

$$\lambda_{2,\text{opt}} = \left[\frac{5(0.08838835)}{64n\rho^2(\nu\pi)^6}\right]^{4/7}$$

which are the values used in the simulation.

Our method of selecting the amount of smoothing for the tests is not practical since it requires knowledge of f. Its advantage is in allowing us to avoid the problems and additional variability that would result from data driven smoothing parameter selection for derivative estimation. With the exception of work by Rice (1986), the problem of bandwidth selection for estimating the derivative of a regression curve has not received the attention of the corresponding problem of smoothing parameter selection for estimation of the regression function alone. The estimators studied in Rice (1986) are derived from an "unbiased" risk type estimator. They are designed for use with tapered Fourier series estimators similar to ours and can be employed to estimate derivatives of any order, provided the regression curve is sufficiently smooth. Under some additional smoothness conditions on f, they would seem to be suitable for use in our setting and we hope to explore this possibility in future work.

In practice, an estimator of $\lambda_{0,\text{opt}}$ will generally be used to fit the data and it may be preferable to use this choice for the smoothing parameter (rather than one designed for derivative estimation) when computing tests. There are many ways to estimate $\lambda_{0,\text{opt}}$ from data, including cross-validation. Thus, one of the goals here was to ascertain, for example, how well the commonly estimated smoothing parameter $\lambda_{0,\text{opt}}$ worked in place of $\lambda_{1,\text{opt}}$ and $\lambda_{2,\text{opt}}$ for T_{1p} and T_{2p} and, more generally, determine the sensitivity of the derivative based tests to suboptimal choices for λ.

Critical values for T_{CK} and T_{mp}, $m = 0,\ 1$ and 2 with nominal level .05 were all obtained using the Buckley/Eagleson approximation discussed in Remark 2 of Section 2. To assess the accuracy of the Buckley/Eagleson approximation for constructing critical values for our tests, we simulated 1000 samples of size 101 under the null model of a constant regression function and then computed the values of our test statistics and their associated critical values for the different choices of the smoothing parameters used in the power study. A different random seed was used for each of the 16 combinations of ρ and ν. The results from this simulation in Table 1 therefore give empirical levels for our tests under various bandwidths. For example, the first entry in the second row of Table 1 gives the proportion of times the null hypothesis of a constant regression function was rejected by T_{0p} when the smoothing parameter was taken to have the value $\lambda_{0,\text{opt}} = [0.265165/32n\rho^2(\nu\pi)^4]^{4/5}$ with $\nu = 1$ and $\rho = .25$.

The proportions in Table 1 are all around 0.05. The only significant (at the .05 level) departures occur for T_{CK} when $\nu = 9$ and $\rho = 1.0$, which significantly exceeds the nominal level, and for the four smallest proportions when $\nu = 1$ where the tests are significantly conservative. We conclude from this that the Buckley/Eagleson approximation performs well and is relatively robust to the amount of smoothing used in the test.

Table 2 contains the empirical powers of our tests against the alternative $f(t) = \rho\cos(2\pi\nu t)$. Each of these entries corresponds to 1000 samples of size 101 generated using the same seeds as for Table 1. Standard errors for the non-unit entries in Table 2 ranged from .0155 to .0009.

Examining Table 2 for when $\nu = 1$, shows all of the tests perform well except for the combination of T_{2p} and $\lambda_{0,\text{opt}}$. For $\nu = 3,\ 9$ or 12, we find that the power

Table 1. Proportion of rejections in 1000 samples of size 101 under null model.

		$\rho = 0.25$	$\rho = 0.5$	$\rho = 1.0$	$\rho = 1.5$
$\nu = 1$	T_{CK}	.041	.047	.050	.044
	T_{0p} $\lambda_{0,\text{opt}}$	.041	.045	.050	.031
	T_{1p} $\lambda_{0,\text{opt}}$	.037	.053	.038	.042
	T_{1p} $\lambda_{1,\text{opt}}$	.040	.043	.035	.035
	T_{2p} $\lambda_{0,\text{opt}}$	.036	.047	.049	.043
	T_{2p} $\lambda_{2,\text{opt}}$	.035	.049	.042	.038
$\nu = 3$	T_{CK}	.040	.051	.051	.057
	T_{0p} $\lambda_{0,\text{opt}}$	.053	.039	.046	.056
	T_{1p} $\lambda_{0,\text{opt}}$	.053	.043	.057	.057
	T_{1p} $\lambda_{1,\text{opt}}$	.043	.038	.052	.054
	T_{2p} $\lambda_{0,\text{opt}}$	.058	.045	.055	.054
	T_{2p} $\lambda_{2,\text{opt}}$	.053	.047	.053	.061
$\nu = 9$	T_{CK}	.050	.045	.067	.055
	T_{0p} $\lambda_{0,\text{opt}}$	.046	.055	.054	.054
	T_{1p} $\lambda_{0,\text{opt}}$	.038	.041	.053	.051
	T_{1p} $\lambda_{1,\text{opt}}$	.042	.045	.053	.048
	T_{2p} $\lambda_{0,\text{opt}}$	.043	.050	.042	.055
	T_{2p} $\lambda_{2,\text{opt}}$	.043	.049	.056	.048
$\nu = 12$	T_{CK}	.053	.043	.058	.053
	T_{0p} $\lambda_{0,\text{opt}}$	.054	.047	.054	.061
	T_{1p} $\lambda_{0,\text{opt}}$	.048	.055	.056	.051
	T_{1p} $\lambda_{1,\text{opt}}$	.043	.053	.054	.059
	T_{2p} $\lambda_{0,\text{opt}}$	.054	.054	.054	.040
	T_{2p} $\lambda_{2,\text{opt}}$	.051	.055	.060	.054

of T_{CK} is approximately equal to the level. This shows T_{CK} is not sensitive to higher frequency alternatives. T_{0p} only has good power for $\nu = 1$ and 3. When $\nu = 9$ and 12, T_{0p} does not have good power but is better than T_{CK}. This is consistent with the results in Eubank and LaRiccia (1993).

In contrast, when the bandwidths are chosen correctly, the test statistics T_{1p} and T_{2p} have good power against all sixteen alternatives. It appears there is little difference in the powers of T_{1p} using either $\lambda_{0,\text{opt}}$ or $\lambda_{1,\text{opt}}$. This is likely a consequence of the fact that $\lambda_{0,\text{opt}}$ and $\lambda_{1,\text{opt}}$ are similar in the sense that both decay to zero at the same rate. Indeed, for this example, we find $\lambda_{0,\text{opt}} \doteq 0.57\lambda_{1,\text{opt}}$. The power of T_{2p} is much more sensitive to the choice of λ than that of T_{1p}. This possibility was anticipated since we know that $\lambda_{0,\text{opt}}$ and $\lambda_{2,\text{opt}}$ converge at different rates.

In summary, there does seem to be some improvement in power using T_{2p} over T_{1p}, although this advantage is offset by the sensitivity of T_{2p} to the choice of λ. Since T_{1p} performs well using $\lambda_{0,\text{opt}}$, and $\lambda_{0,\text{opt}}$ can be readily estimated from the data, the test based on the first derivative may be more effective in practice.

Table 2. Proportion of rejections in 1000 samples of size 101.

		$\rho = 0.25$	$\rho = 0.5$	$\rho = 1.0$	$\rho = 1.5$
$\nu = 1$	T_{CK}	.319	.899	1.000	1.000
	T_{0p} $\lambda_{0,\mathrm{opt}}$	.319	.893	1.000	1.000
	T_{1p} $\lambda_{0,\mathrm{opt}}$	.308	.801	1.000	1.000
	T_{1p} $\lambda_{1,\mathrm{opt}}$	.314	.867	1.000	1.000
	T_{2p} $\lambda_{0,\mathrm{opt}}$	.199	.254	.324	.339
	T_{2p} $\lambda_{2,\mathrm{opt}}$	.307	.844	1.000	1.000
$\nu = 3$	T_{CK}	.040	.056	.074	.106
	T_{0p} $\lambda_{0,\mathrm{opt}}$	.117	.695	1.000	1.000
	T_{1p} $\lambda_{0,\mathrm{opt}}$	.256	.759	.999	1.000
	T_{1p} $\lambda_{1,\mathrm{opt}}$	.241	.806	1.000	1.000
	T_{2p} $\lambda_{0,\mathrm{opt}}$	.214	.404	.618	.741
	T_{2p} $\lambda_{2,\mathrm{opt}}$	.270	.776	.999	1.000
$\nu = 9$	T_{CK}	.050	.045	.067	.056
	T_{0p} $\lambda_{0,\mathrm{opt}}$	.059	.290	.998	1.000
	T_{1p} $\lambda_{0,\mathrm{opt}}$	.120	.610	.999	1.000
	T_{1p} $\lambda_{1,\mathrm{opt}}$	.086	.602	.999	1.000
	T_{2p} $\lambda_{0,\mathrm{opt}}$	.170	.452	.840	.970
	T_{2p} $\lambda_{2,\mathrm{opt}}$	.152	.625	.999	1.000
$\nu = 12$	T_{CK}	.053	.043	.059	.053
	T_{0p} $\lambda_{0,\mathrm{opt}}$	.063	.225	.987	1.000
	T_{1p} $\lambda_{0,\mathrm{opt}}$	.105	.557	1.000	1.000
	T_{1p} $\lambda_{1,\mathrm{opt}}$	.082	.502	1.000	1.000
	T_{2p} $\lambda_{0,\mathrm{opt}}$	.152	.465	.877	.974
	T_{2p} $\lambda_{2,\mathrm{opt}}$	.136	.575	1.000	1.000

Acknowledgements

The author is very grateful to Professor R. L. Eubank of Texas A&M University, College Station, Texas, USA, for reading this manuscript, helpful suggestions and valuable comments. The author would also liked to thank the two referees for their careful comments which lead to improvements in the presentation of the results.

Appendix

Define $\boldsymbol{y}_n = (y_1, \ldots, y_n)'$, $\boldsymbol{f}_n = (f(t_{1n}), \ldots, f(t_{nn}))'$, $\boldsymbol{g}_n = (g(t_{1n}), \ldots, g(t_{nn}))'$ and $X_{nn} = [x_j(t_{kn})]_{k=1,\ldots,n;|j|\leq(n-1)/2}$, where $x_j(t) = e^{2\pi ijt}$ for $e^{it} = \cos t + i \sin t$ and $i^2 = -1$. Since $n^{-1}\sum_{k=1}^{n} e^{-2\pi ij(k-1)/n} e^{2\pi il(k-1)/n} = 1$, if $j = l$, $= 0$, if $j \neq l$, we have $X_{nn}^* X_{nn} = nI_n$, where the $*$ notation is used to indicate the complex conjugate transpose of a matrix and I_n is the $n \times n$ identity matrix. The following notation will also be needed for the proofs

$$\widetilde{\sum_{l,j}} = \sum_{\substack{l \\ l\neq j}} \sum_{j}, \quad \widetilde{\sum_{l,j,k}} = \sum_{\substack{l \\ l\neq j}} \sum_{\substack{j \\ j\neq k}} \sum_{\substack{k \\ k\neq l}} \quad \text{and} \quad \widetilde{\sum_{i,j,k,l}} = \sum_{\substack{i \\ i\neq j \\ j\neq l}} \sum_{\substack{j \\ j\neq k \\ k\neq l}} \sum_{\substack{k \\ i\neq l}} \sum_{\substack{l \\ j\neq k}}$$

The proof of Theorems 2.1, 2.2 and 2.3 requires a result from Eubank and LaRiccia (1993). We state this formally below as Lemma A.1. We also need a result concerning the asymptotic distribution of quadratic forms provided in Lemma A.2. Its proof is a simple application of results in de Jong (1987). See Jayasuriya (1990) for details.

LEMMA A.1. (Eubank and LaRiccia (1993)). *Let* $g \in W^2_{2,\mathrm{per}}[0,1]$, *define* $b_{jn} = n^{-1}\sum_{k=1}^n g(t_{kn})e^{-2\pi ijt_{kn}}$ *and set* $b_j = \int_0^1 g(t)e^{-2\pi ijt}dt$. *Then* $|b_{jn} - b_j| = o(n^{-2})$, *uniformly in* $|j| \le (n-1)/2$.

LEMMA A.2. *Let* $\boldsymbol{y}_n = (y_1, \ldots, y_n)'$ *be a random vector and* $\boldsymbol{f}_n = (f_{1n}, \ldots, f_{nn})' \in \boldsymbol{R}^n$. *Define* $\boldsymbol{\epsilon}_n = (\epsilon_1, \ldots, \epsilon_n)' = \boldsymbol{y}_n - \boldsymbol{f}_n$ *and suppose* $\epsilon_1, \ldots, \epsilon_n$ *are independent, identically distributed random variables with* $E[\epsilon_1] = 0$, $\mathrm{Var}[\epsilon_1] = \sigma^2$ *and* $0 < E[\epsilon_1^4] < \infty$. *Let* M_n *be a symmetric* $n \times n$ *matrix of constants and* m_{nlj} *be its* lj*-th element with* $m_{nlj}^{(k)}$ *denoting the* lj*-th element of* M_n^k, *for* $k = 2, 3, \ldots$. *Define* $\sigma^2(n) = \sum_{j=1}^n (m_{njj}^{(2)} - m_{njj}^2)$,

$$\alpha_1 = \widetilde{\sum_{l,j}} m_{nlj}^4,$$

$$\alpha_2 = \widetilde{\sum_{l,j,k}} m_{nlj}^2 m_{nlk}^2 \quad \textit{and}$$

$$\alpha_3 = \widetilde{\sum_{i,j,k,l}} m_{nij}m_{nik}m_{nlj}m_{nlk}.$$

Then

$$A_n = \frac{\boldsymbol{y}_n' M_n \boldsymbol{y}_n - \sigma^2 \operatorname{tr} M_n - \boldsymbol{f}_n' M_n \boldsymbol{f}_n}{\sigma^2\sqrt{2\operatorname{tr} M_n^2}}$$

converges in distribution to a standard normal random variable as $n \to \infty$ *if,*

(A) $\dfrac{\sum_j m_{njj}^2}{\operatorname{tr} M_n^2} \to 0 \quad$ *as* $\quad n \to \infty$,

(B) $\dfrac{\boldsymbol{f}_n' M_n^2 \boldsymbol{f}_n}{\operatorname{tr} M_n^2} \to 0 \quad$ *as* $\quad n \to \infty \quad$ *and*

(C) $\alpha_j = o(\sigma^4(n)) \quad$ *for* $j = 1, 2, 3 \quad$ *as* $\quad n \to \infty$.

PROOF OF THEOREM 2.1. Eubank and LaRiccia (1993) prove Conditions (A) and (B) are true for T_{0p}. Thus, it suffices to show that Condition (C) of

Lemma A.2 is satisfied. We have $T_{0p} = \sum'_{|j|\le(n-1)/2} |\tilde{a}_{jn}|^2/(1+\lambda(2\pi j)^4)^2 = n^{-2}\boldsymbol{y}_n' X H_n X^* \boldsymbol{y}_n$. So, we can apply Lemma A.2 using $M_n = n^{-2}XH_nX^*$ with H_n an $n\times n$ diagonal matrix with j-th diagonal element $h_{jj} = 1/(1+\lambda(2\pi j)^4)^2$, $|j| \le (n-1)/2$, $j \ne 0$ and $h_{00} = 0$.

We have $m_{nlk} = n^{-2}\sum'_{|j|\le(n-1)/2} 1/(1+\lambda(2\pi j)^4)^2 e^{-2\pi i j((l-1)/n)} \cdot e^{2\pi i j((k-1)/n)}$. Therefore, $|m_{nlk}| \le n^{-2}\sum'_{|j|\le(n-1)/2} 1/(1+\lambda(2\pi j)^4)^2$. Now observe that

$$\sum_{|j|\le(n-1)/2}' \frac{1}{(1+\lambda(2\pi j)^4)^2} = 2\int_0^{(n-1)/2} \frac{1}{(1+\lambda(2\pi y)^4)^2}dy + O(1)$$
$$\sim 2\lambda^{-1/4}\int_0^\infty \frac{1}{(1+(2\pi x)^4)^2}dx.$$

Thus, $m_{nlk} = O(n^{-2}\lambda^{-1/4})$.

Since $M_n^2 = n^{-3}XH_n^2X^*$, the eigenvalues of M_n^2 are

$$\text{(A.1)} \qquad E_{0j} = \begin{cases} 0, & \text{if } j = 0, \\ n^{-2}\dfrac{1}{(1+\lambda(2\pi j)^4)^4}, & \text{if } |j| \le \dfrac{n-1}{2} \text{ and } j \ne 0. \end{cases}$$

Using, $\sum'_{|j|\le(n-1)/2} 1/(1+\lambda(2\pi j)^4)^4 \sim 2\lambda^{-1/4}\int_0^\infty 1/(1+(2\pi x)^4)^4 dx = C_0\lambda^{-1/4}$ with $C_0 = \pi^{-1}\int_0^\infty 1/(1+x^4)^4 dx$, we then obtain $\operatorname{tr} M_n^2 = n^{-2}\sum'_{|j|\le(n-1)/2} 1/(1+\lambda(2\pi j)^4)^4 \sim C_0 n^{-2}\lambda^{-1/4} = O(n^{-2}\lambda^{-1/4})$.

To show that Condition (C) holds, we need to prove that $\alpha_j = o(\sigma^4(n))$ for $j = 1,2,3$ in Lemma A.2. We have $\sigma^2(n) = \sum'_{|j|\le(n-1)/2}(m_{njj}^{(2)} - m_{njj}^2) = \operatorname{tr} M_n^2 - \sum'_{|j|\le(n-1)/2} m_{njj}^2$. From this and Condition (A) we see that $\sigma^2(n)/\operatorname{tr} M_n^2 = 1 - (\sum'_{|j|\le(n-1)/2} m_{njj}^2)/\operatorname{tr} M_n^2 \to 1$ as $n\to\infty$, $\lambda\to 0$ and $n\lambda^{1/2}\to\infty$. Therefore, $\sigma^2(n) \sim \operatorname{tr} M_n^2$ as $n\to\infty$, $\lambda\to 0$ and $n\lambda^{1/2}\to\infty$ and $\sigma^4(n) \sim C_0^2 n^{-4}\lambda^{-1/2}$.

First consider $\alpha_1 = \tilde{\sum}_{l,j} m_{nlj}^4 = n^2 O(n^{-2}\lambda^{-1/4})^4 = O(n^{-6}\lambda^{-1})$. We have $\alpha_1/\sigma^4(n) = O(n^{-6}\lambda^{-1})/(C_0^2 n^{-4}\lambda^{-1/2} + o(\sigma^4(n))) = O(n^{-2}\lambda^{-1/2}) = o(1)$, as $n\to\infty$, $\lambda\to 0$ and $n\lambda^{1/2}\to\infty$. Similarly, $\alpha_2 = \tilde{\sum}_{l,j,k} m_{nlj}^2 m_{nlk}^2 = n^3 O(n^{-4}\lambda^{-1/2}) \cdot O(n^{-4}\lambda^{-1/2}) = O(n^{-5}\lambda^{-1})$ and $\alpha_2/\sigma^4(n) = O(n^{-5}\lambda^{-1})/(C_0^2 n^{-4}\lambda^{-1/2} + o(\sigma^4(n))) = O(n^{-1}\lambda^{-1/2}) = o(1)$.

Finally,

$$\alpha_3 = \widetilde{\sum_{i,j,k,l}} m_{nij}m_{nik}m_{nlj}m_{nlk} = \widetilde{\sum_{i,j,k}} m_{nij}m_{nik} \sum_{\substack{l \\ l\ne i,j,k}} m_{nlj}m_{nlk}$$
$$= \widetilde{\sum_{i,j,k}} m_{nij}m_{nik}(m_{njk}^{(2)} - m_{nij}m_{nik} - m_{njj}m_{njk} - m_{nkj}m_{nkk})$$
$$= \widetilde{\sum_{j,k}}(m_{njk}^{(2)} - m_{njj}m_{njk} - m_{nkj}m_{nkk}) \sum_{\substack{i \\ i\ne j, i\ne k}} m_{nij}m_{nik} - \widetilde{\sum_{i,j,k}} m_{nij}^2 m_{nik}^2$$

$$= \widetilde{\sum_{j,k}}(m_{njk}^{(2)} - m_{njj}m_{njk} - m_{nkj}m_{nkk})^2 + \alpha_2$$

$$= \widetilde{\sum_{j,k}}(m_{njk}^{(2)})^2 + \widetilde{\sum_{j,k}} m_{njj}^2 m_{njk}^2 - \widetilde{\sum_{j,k}} m_{njk}^{(2)} m_{njj} m_{njk}$$

$$+ \widetilde{\sum_{j,k}} m_{njj} m_{nkk} m_{njk}^2 - \alpha_2.$$

We will deal with each term in this sum separately.

First, $\tilde{\sum}_{j,k}(m_{njk}^{(2)})^2 = \sum'_{|j|\le(n-1)/2} m_{njj}^{(4)} - \sum'_{|j|\le(n-1)/2}(m_{njj}^{(2)})^2 \le \sum'_{|j|\le(n-1)/2} m_{njj}^{(4)}$. By definition, $\operatorname{tr} M_n^4 = \sum'_{|j|\le(n-1)/2} m_{njj}^{(4)}$ with $M_n^4 = n^{-5} X H_n^4 X^*$ and $|m_{nlk}^{(4)}| \le n^{-5}\sum'_{|j|\le(n-1)/2} 1/(1+\lambda(2\pi j)^4)^8 \sim 2n^{-5}\lambda^{-1/4} \cdot \int_0^\infty 1/(1+(2\pi x)^4)^8 dx$. So, $m_{nlk}^{(4)} = O(n^{-5}\lambda^{-1/4})$ and $\sum'_{|j|\le(n-1)/2} m_{njj}^{(4)} = O(n^{-4}\lambda^{-1/4})$. Also, we have $m_{nlk}^2 = O(n^{-4}\lambda^{-1/2})$ which gives $\tilde{\sum}_{j,k} m_{njj}^2 m_{njk}^2 = O(n^{-6}\lambda^{-1})$.

Next consider $\tilde{\sum}_{j,k} m_{njk}^{(2)} m_{njj} m_{njk}$. Since $|m_{nlk}^{(2)}| \le n^{-3}\sum'_{|j|\le(n-1)/2} 1/(1+\lambda(2\pi j)^4)^4 = O(n^{-3}\lambda^{-1/4})$, we have

$$\left|\widetilde{\sum_{j,k}} m_{njk}^{(2)} m_{njj} m_{njk}\right| = n^2 O(n^{-3}\lambda^{-1/4}) O(n^{-2}\lambda^{-1/4}) O(n^{-2}\lambda^{-1/4})$$
$$= O(n^{-5}\lambda^{-3/4})$$

and $\tilde{\sum}_{j,k} m_{njj} m_{nkk} m_{njk}^2 = n^2 O(n^{-2}\lambda^{-1/4}) O(n^{-2}\lambda^{-1/4}) O(n^{-4}\lambda^{-1/2}) = O(n^{-6}\lambda^{-1})$. Therefore,

$$\frac{\alpha_3}{\sigma^4(n)} = \frac{O(n^{-4}\lambda^{-1/4})}{C_0^2 n^{-4}\lambda^{-1/2} + o(\sigma^4(n))} + \frac{O(n^{-6}\lambda^{-1})}{C_0^2 n^{-4}\lambda^{-1/2} + o(\sigma^4(n))}$$
$$+ \frac{O(n^{-5}\lambda^{-3/4})}{C_0^2 n^{-4}\lambda^{-1/2} + o(\sigma^4(n))} + \frac{O(n^{-6}\lambda^{-1})}{C_0^2 n^{-4}\lambda^{-1/2} + o(\sigma^4(n))} + \frac{\alpha_2}{\sigma^4(n)}$$
$$= O(\lambda^{1/4}) + O(n^{-2}\lambda^{-1/2}) + O(n^{-1}\lambda^{-1/4}) + O(n^{-2}\lambda^{-1/2}) + o(1)$$
$$= o(1).$$

Consequently, $\alpha_3 = o(\sigma^4(n))$ as $n \to \infty$, $\lambda \to 0$ and $n\lambda^{1/2} \to \infty$ and Condition (C) holds.

We now apply Lemma A.2 to get $Z_{0p} - (\boldsymbol{f}_n' M_n \boldsymbol{f}_n / \sigma^2\sqrt{2\operatorname{tr} M_n^2}) \xrightarrow{\mathcal{D}} N(0,1)$ as $n \to \infty$, $\lambda \to 0$ and $n\lambda^{1/2} \to \infty$. But, $\boldsymbol{f}_n = h(n)\boldsymbol{g}_n$ for $\boldsymbol{g}_n = (g(t_{1n}), \ldots, g(t_{nn}))'$ and $h(n) = n^{-1/2}\lambda^{-1/16}$. Therefore,

$$\frac{\boldsymbol{f}_n' M_n \boldsymbol{f}_n}{\sigma^2\sqrt{2\operatorname{tr} M_n^2}} \sim \frac{n^{-1}\lambda^{-1/8}\boldsymbol{g}_n' M_n \boldsymbol{g}_n}{\sqrt{2C_0}\sigma^2 n^{-1}\lambda^{-1/8}} = \frac{\boldsymbol{g}_n' M_n \boldsymbol{g}_n}{\sigma^2\sqrt{2C_0}} \to \frac{\|g\|^2}{\sigma^2\sqrt{2C_0}}$$

as $n \to \infty$, $\lambda \to 0$ and $n\lambda^{1/2} \to \infty$, because Eubank and LaRiccia (1993) show that $g_n' M_n g_n = \sum'_{|j|\le(n-1)/2} |b_{jn}|^2/(1+\lambda(2\pi j)^4)^2 \to \|g\|^2$, where the b_{jn} are defined in Lemma A.1. The proof is completed by an application of Slutsky's Theorem. □

PROOF OF THEOREM 2.2. The proof of Theorem 2.2 is similar to that for Theorem 2.1. It also uses Lemma A.2 except that now M_n^2 is a matrix with eigenvalues

$$E_{1j} = \begin{cases} 0, & \text{if } j = 0, \\ n^{-2}\dfrac{(2\pi j)^4}{(1+\lambda(2\pi j)^4)^4}, & \text{if } |j| \le \dfrac{n-1}{2} \text{ and } j \ne 0. \end{cases} \tag{A.2}$$

The details can be found in Chen (1992). □

PROOF OF THEOREM 2.3. We have $T_{2p} = \sum'_{|j|\le(n-1)/2}((2\pi j)^4|\tilde{a}_{jn}|^2)/(1+\lambda(2\pi j)^4)^2 = n^{-2}y_n' X H_n X^* y_n$, so we apply Lemma A.2 with $M_n = n^{-2}XH_nX^*$ and H_n a diagonal matrix with elements $h_{00} = 0$ and $h_{jj} = (2\pi j)^4/(1+\lambda(2\pi j)^4)^2$ for $|j| \le (n-1)/2$ and $j \ne 0$.

First note that $m_{nlk} = n^{-2}\sum'_{|j|\le(n-1)/2}(2\pi j)^4/(1+\lambda(2\pi j)^4)^2 e^{-2\pi ij(l-1)/n} \cdot e^{2\pi ij(k-1)/n}$ and $|m_{nlk}| \le n^{-2}\sum'_{|j|\le(n-1)/2}(2\pi j)^4/(1+\lambda(2\pi j)^4)^2$. An integral approximation then gives

$$\sum_{|j|\le(n-1)/2}{}' \frac{(2\pi j)^4}{(1+\lambda(2\pi j)^4)^2} \sim 2\lambda^{-5/4}\int_0^\infty \frac{(2\pi x)^4}{(1+(2\pi x)^4)^2}dx.$$

Thus, $m_{nlk} = O(n^{-2}\lambda^{-5/4})$.

The eigenvalues of M_n^2 are

$$E_{2j} = \begin{cases} 0, & \text{if } j = 0, \\ n^{-2}\dfrac{(2\pi j)^8}{(1+\lambda(2\pi j)^4)^4}, & \text{if } |j| \le \dfrac{n-1}{2} \text{ and } j \ne 0. \end{cases} \tag{A.3}$$

Thus,

$$\begin{aligned}\operatorname{tr} M_n^2 &= n^{-2}\sum_{|j|\le(n-1)/2}{}' \frac{(2\pi j)^8}{(1+\lambda(2\pi j)^4)^4} \sim 2n^{-2}\lambda^{-9/4}\int_0^\infty \frac{(2\pi x)^8}{(1+(2\pi x)^4)^4}dx \\ &= O(n^{-2}\lambda^{-9/4}).\end{aligned}$$

Consequently, Condition (A) of Lemma A.2 holds because

$$\begin{aligned}\left(\sum_{|j|\le(n-1)/2}{}' m_{njj}^2\right)\Big/\operatorname{tr} M_n^2 &\sim (nO(n^{-4}\lambda^{-5/2}))/(C_2 n^{-2}\lambda^{-9/4}) \\ &= O(n^{-1}\lambda^{-1/4}) = o(1)\end{aligned}$$

as $n \to \infty$, $\lambda \to 0$ and $n\lambda^{9/8} \to \infty$.

For Condition (B) note that, since $h(n) = n^{-1/2}\lambda^{-9/16}$, $\boldsymbol{f}_n' M_n^2 \boldsymbol{f}_n = h(n)^2 \boldsymbol{g}_n' \cdot M_n^2 \boldsymbol{g}_n = n^{-2}\lambda^{-9/8} \sum'_{|j| \le (n-1)/2} ((2\pi j)^8 |b_{jn}|^2)/(1 + \lambda(2\pi j)^4)^4 \le n^{-2}\lambda^{-9/8}\lambda^{-1} \cdot \sum'_{|j| \le (n-1)/2} ((2\pi j)^4 |b_{jn}|^2)/(1 + \lambda(2\pi j)^4)^2$. Thus, to verify the condition we need only show that $\sum'_{|j| \le (n-1)/2} ((2\pi j)^4 |b_{jn}|^2)/(1 + \lambda(2\pi j)^4)^2 \to \|g''\|^2$. For this purpose observe that $\|g''\|^2 = \int_0^1 |g''|^2 dt = \sum_{j=-\infty}^{\infty} |b_j''|^2 = \sum_{j=-\infty}^{\infty} (2\pi j)^4 |b_j|^2 = \sum'_{|j| \le (n-1)/2} (2\pi j)^4 |b_j|^2 + o(1)$. Thus,

$$
\begin{aligned}
&\left| \sum_{|j| \le (n-1)/2}' \frac{(2\pi j)^4 |b_{jn}|^2}{(1 + \lambda(2\pi j)^4)^2} - \sum_{|j| \le (n-1)/2}' (2\pi j)^4 |b_j|^2 \right| \\
&\quad \sim \left| \sum_{|j| \le (n-1)/2}' \frac{(2\pi j)^4 j^{-(5+\delta)}}{(1 + \lambda(2\pi j)^4)^2} - \sum_{|j| \le (n-1)/2}' (2\pi j)^4 j^{-(5+\delta)} \right| \\
&\quad = \left| \sum_{|j| \le (n-1)/2}' \frac{(2\pi)^4 j^{-(1+\delta)}}{(1 + \lambda(2\pi j)^4)^2} - \sum_{|j| \le (n-1)/2}' (2\pi)^4 j^{-(1+\delta)} \right| \\
&\quad = (2\pi)^4 \sum_{|j| \le (n-1)/2}' \left| \frac{j^{-(1+\delta)}(1 - (1 + \lambda(2\pi j)^4)^2)}{(1 + \lambda(2\pi j)^4)^2} \right| \\
&\quad \le \sum_{|j| \le (n-1)/2}' \frac{2\lambda(2\pi)^8 j^{3-\delta}}{(1 + \lambda(2\pi j)^4)^2} + \sum_{|j| \le (n-1)/2}' \frac{\lambda^2 (2\pi)^{12} j^{7-\delta}}{(1 + \lambda(2\pi j)^4)^2}.
\end{aligned}
$$

By integral approximations

$$
\sum_{|j| \le (n-1)/2}' \frac{2\lambda(2\pi)^8 j^{3-\delta}}{(1 + \lambda(2\pi j)^4)^2} \sim 4\lambda^{\delta/4}(2\pi)^{4+\delta} \int_0^\infty \frac{y^{3-\delta}}{(1 + y^4)^2} dy = O(\lambda^{\delta/4})
$$

and

$$
\sum_{|j| \le (n-1)/2}' \frac{\lambda^2 (2\pi)^{12} j^{7-\delta}}{(1 + \lambda(2\pi j)^4)^2} \sim 2\lambda^{\delta/4}(2\pi)^{4+\delta} \int_0^1 \frac{y^{7-\delta}}{(1 + (2\pi)^4 y^4)^2} dy = O(\lambda^{\delta/4}).
$$

Therefore, $|\sum'_{|j| \le (n-1)/2} ((2\pi j)^4 |b_{jn}|^2)/(1 + \lambda(2\pi j)^4)^2 - \|g''\|^2| \le O(\lambda^{\delta/4}) + O(\lambda^{\delta/4}) + o(1) = o(1)$ and $(\boldsymbol{f}_n' M_n^2 \boldsymbol{f}_n)/\operatorname{tr} M_n^2 = O(\lambda^{1/8})$.

The proof that Condition (C) holds follows along similar lines as in the proof of Theorem 2.1. We omit the details.

Lemma A.2 now gives $Z_{2p} - (\boldsymbol{f}_n' M_n \boldsymbol{f}_n / \sigma^2 \sqrt{2 \operatorname{tr} M_n^2}) \xrightarrow{\mathcal{D}} N(0,1)$ as $n \to \infty$, $\lambda \to 0$ and $n\lambda^{9/8} \to \infty$, for $h(n) = n^{-1/2}\lambda^{-9/16}$. The result then follows from observing that

$$
\frac{\boldsymbol{f}_n' M_n \boldsymbol{f}_n}{\sigma^2 \sqrt{2 \operatorname{tr} M_n^2}} = \frac{n^{-1}\lambda^{-9/8} \boldsymbol{g}_n' M_n \boldsymbol{g}_n}{\sqrt{2C_2}\sigma^2 n^{-1}\lambda^{-9/8}} = \frac{\boldsymbol{g}_n' M_n \boldsymbol{g}_n}{\sigma^2 \sqrt{2C_2}} \to \frac{\|g''\|^2}{\sigma^2 \sqrt{2C_2}}
$$

as $n \to \infty$, $\lambda \to 0$ and $n\lambda^{9/8} \to \infty$, since $\boldsymbol{g}_n' M_n \boldsymbol{g}_n \to \|g''\|^2$ if $|b_j|^2 \sim j^{-(5+\delta)}$ for $\delta > 1$. $\square$

References

Buckley, M. J. (1991). Detecting a smooth signal: optimality of cusum based procedures, *Biometrika*, **78**, 253–262.

Buckley, M. J. and Eagleson, G. K. (1988). An approximation to the distribution of quadratic forms in normal random variables, *Austral. J. Statist.*, **30**, 149–163.

Chen, J. C. (1992). Testing goodness of fit of polynomial models via spline smoothing techniques, Ph.D. Dissertation, Texas A&M University (unpublished).

Cox, D. D. and Koh, E. (1989). A smoothing spline based test of model adequacy in polynomial regression, *Ann. Inst. Statist. Math.*, **41**, 383–400.

Cox, D. D., Koh, E., Wahba, G. and Yandell, B. (1988). Testing the (parametric) null model hypothesis in (semiparametric) partial and generalized spline models, *Ann. Statist.*, **16**, 113–119.

De Jong, P. (1987). A central limit theorem for generalized quadratic forms, *Probab. Theory Related Fields*, **25**, 261–277.

Eubank, R. L. (1988). *Spline Smoothing and Nonparametric Regression*, Marcel Dekker, New York.

Eubank, R. L. and Hart, J. D. (1992). Testing goodness-of-fit in regression via order selection criteria, *Ann. Statist.*, **20**, 1412–1425.

Eubank, R. L. and Hart, J. D. (1993). Commonality of cusum, von Neumann and smoothing-based goodness-of-fit tests, *Biometrika*, **80**, 89–98.

Eubank, R. L. and LaRiccia, V. N. (1993). Testing for no effect in nonparametric regression, *J. Statist. Plann. Inference*, **36**, 1–14.

Eubank, R. L. and Spiegelman, C. H. (1990). Testing the goodness of fit of a linear model via nonparametric regression techniques, *J. Amer. Statist. Assoc.*, **85**, 387–392.

Gasser, T., Sroka, L. and Jennen-Steinmetz, C. (1986). Residual variance and residual pattern in nonlinear regression, *Biometrika*, **73**, 625–633.

Graybill, F. A. (1976). *Theory and Application of the Linear Model*, Duxbury Press, Boston.

Hall, P., Kay, J. W. and Titterington, D. M. (1990). Asymptotically optimal difference-based estimation of variance in nonparametric regression, *Biometrika*, **77**, 521–528.

Härdle, W. and Mammen, E. (1988). Comparing nonparametric versus parametric regression fits (unpublished).

Hart, J. D. and Wehrly, T. E. (1992). Kernel regression when the boundary region is large, with an application to testing the adequacy of polynomial models. *J. Amer. Statist. Assoc.*, **87**, 1018–1024.

Jayasuriya, B. R. (1990). Testing for polynomial regression using nonparametric regression techniques, Ph.D. Dissertation, Texas A&M University.

Kleinbaum, D. G., Kupper, L. L. and Muller, K. E. (1988). *Application Regression Analysis and Other Multivariable Methods*, PWS-KENT Publishing, Boston.

Müller, H. G. (1992). Goodness-of-fit diagnostics for regression models, *Scand. J. Statist.*, **19**, 157–172.

Munson, P. J. and Jernigan, R. W. (1989). A cubic spline extension of the Durbin-Watson test, *Biometrika*, **76**, 39–47.

Raz, J. (1990). Testing for no effect when estimating a smooth function by nonparametric regression: a randomization approach, *J. Amer. Statist. Assoc.*, **85**, 132–138.

Rice, J. (1986). Bandwidth choice for differentiation, *J. Multivariate Anal.*, **19**, 251–264.

Rice, J. and Rosenblatt, M. (1981). Integrated mean squared error of a smoothing spline, *J. Approx. Theory*, **33**, 353–369.

Staniswalis, J. G. and Severini, T. A. (1991). Diagnostics for assessing regression models, *J. Amer. Statist. Assoc.*, **86**, 684–692.

von Neumann, J. (1941). Distribution of the ratio of the mean squared successive difference to the variance, *Ann. Math. Statist.*, **12**, 367–395.