Testing Parametric Assumptions on Band- or Time-Limited Signals Under Noise

Nicolai Bissantz, Hajo Holzmann, and Axel Munk

***Abstract*—This paper considers the problem of testing parametric assumptions on signals f from which only noisy observations $y_k = f(\tau k) + \epsilon_k$ are available, and where the signal is assumed to be either band-limited or time-limited. To this end, the signal is reconstructed by an estimator based on the Whittaker–Shannon (WS) sampling theorem with oversampling. As test statistic, the minimal L_2 distance between the estimated signal and the parametric model is used. To construct appropriate tests, the asymptotic distribution of the test statistic is derived both under the hypothesis of the validity of the parametric model and under fixed local alternatives. As a byproduct, the asymptotic distribution of the integrated square error of the estimator is computed, which is of interest by itself, e.g., for the analysis of a cross-validated bandwidth selector.**

***Index Terms*—Asymptotic normality, band-limited signals, goodness of fit, non-band-limited signals, oversampling, rate of convergence, signal recovery, Whittaker–Shannon (WS) sampling theorem.**

I. Introduction

THE problem of reconstructing a nonparametric signal f from data which is corrupted by random noise, has been investigated intensively in recent years both in statistics [10]–[12]and in engineering [21], [31]. In a number of applications in communication theory (e.g., [35]), the signal f is a function of time t and assumed to be in the class of band-limited signals, i.e., signals for which the Fourier transform has compact support. Band-limited functions also appear e.g., as point spread functions in optics or autocorrelation functions in crystallography [20]. Throughout the following, we define the Fourier transform of a signal $f \in L_2(\mathbb{R})$ as $F = \mathcal{F}(f)(\omega) = \int_{\mathbb{R}} f(t)e^{-it\omega}dt$, and write $f \in \mathrm{BL}(\tilde{\Omega})$ if the support of the Fourier transform of f is contained in $[-\tilde{\Omega}, \tilde{\Omega}]$. From the Paley–Wiener theorem, band-limited signals extend to entire functions on the complex plain, and thus can never have compact support. It is well known that they can be recovered from a countable number of samples, i.e., if $f \in \mathrm{BL}(\tilde{\Omega})$ and $\tau \leq \pi/\tilde{\Omega}$, then

$$f(t) = \sum_{k\in\mathbb{Z}} f(\tau k)\mathrm{sinc}(\pi/\tau(t-\tau k)). \qquad (1)$$

Manuscript received August 2, 2004; revised January 13, 2005. The work of H. Holzmann and A. Munk was supported by the DFG under Grant MU 1230/8-1.

The authors are with the Institut für Mathematische Stochastik, University of Göttingen, 37073 Göttingen, Germany (e-mail: bissantz@math.uni-goettingen.de; holzmann@math.uni-goettingen.de; munk@math.uni-goettingen.de).

Communicated by P. L. Bartlett, Associate Editor for Pattern Recognition, Statistical Learning and Inference.
Digital Object Identifier 10.1109/TIT.2005.856957

Here $\mathrm{sinc}(x) = \sin(x)/x$ and $\mathrm{sinc}(0) = 1$. The expansion (1) is called the Whittaker–Shannon (WS) sampling theorem or simply the cardinal expansion of f; convergence in (1) is uniform on bounded intervals. We refer to [4], [16], [24], [25], [38] for further information on the WS sampling theorem.

The aim of this paper is two-fold. First, we derive the distributional limit of the integrated square error of an estimator of f, which was introduced by Pawlak and Stadtmüller [31] and second, this will be used to check parametric assumptions on f. The simplest case is to test whether there is a signal at all, i.e., $f \equiv 0$ (see [34]). Following [31] we introduce the model

$$y_k = f(\tau k) + \epsilon_k, \qquad k \in \mathbb{Z},\ \tau > 0 \qquad (2)$$

where we observe a finite number of y_k; $|k| \leq n$. Here, $(\epsilon_k)_{k\in\mathbb{Z}}$ is an independent and identically distributed (i.i.d.) noise process with $E\epsilon_k = 0$, $E\epsilon_k^4 < \infty$. We set $\sigma^2 = E\epsilon_k^2$. If the signal f is band-limited with $f \in \mathrm{BL}(\tilde{\Omega})$ and if $\tau \leq \pi/\tilde{\Omega}$, a first natural possibility for estimating f, based on the cardinal expansion, is given as

$$\tilde{f}_n(t) = \sum_{|k|\leq n} y_k \mathrm{sinc}(\pi/\tau(t-\tau k)).$$

Although this estimator is evidently asymptotically unbiased, its asymptotic variance is equal to that of the original observation since it interpolates noise, see [24] or [33]. In order to obtain a consistent estimator, the method of oversampling can be used. Recall that the cardinal series expansion with oversampling is given by

$$f(t) = \tau \sum_{k\in\mathbb{Z}} f(\tau k)\frac{\sin(\Omega(t-\tau k))}{\pi(t-\tau k)}, \qquad \Omega \geq \tilde{\Omega},\ \tau \leq \pi/\Omega \qquad (3)$$

where $\Omega\tau/\pi$ corresponds to the sampling rate. Convergence in (3) is again uniform on bounded intervals. Based on the expansion (3), the estimator of $f(t)$ is given by

$$\hat{f}_n(t) = \tau \sum_{|k|\leq n} y_k \frac{\sin(\Omega(t-\tau k))}{\pi(t-\tau k)}, \qquad \tau \leq \pi/\Omega \qquad (4)$$

cf. [24], [31]–[33]. It can be shown for band-limited signals $f(t)$ that $\hat{f}_n(t)$ is pointwise consistent if $\tau \to 0$ and $n\tau \to \infty$, i.e., $\hat{f}_n(t) \to f(t)$. Moreover, under certain assumptions on the tail behavior of f, estimates on the mean integrated square error (MISE)

$$\mathrm{MISE}(\hat{f}_n) = E\int_{\mathbb{R}} (\hat{f}_n(t) - f(t))^2 dt$$

are derived in [30], [32]. It has been further shown there that for band-limited signals in model (2), the rate of decay of the MISE of the estimator $\hat{f}_n$ is reasonably fast.

In this paper, the distributional limit of the integrated square error (ISE) of the estimator $\hat{f}_n$

$$\mathrm{ISE}(\hat{f}_n) = \int_{\mathbb{R}} (\hat{f}_n(t) - f(t))^2 dt$$

is investigated. This is a classical theme in the statistical literature, because the asymptotics of the ISE yields information on the variability of cross-validation when used as an automatic selector of the smoothing parameter τ. In fact, asymptotic normality of the ISE was first obtained for kernel density estimators (see [2], [13]) and for kernel regression estimators in a random design ([14], [23]). For a regression model on a compact interval with fixed design, asymptotic normality of the ISE was proved in [18] by arguing via a central limit theorem for martingales. However, all these results cannot be applied to the model (2). First, in regression models, theory for an unbounded time domain has not been developed, and second, the sinc kernel $K(t) = \sin(t)/(\pi t)$ is not integrable. In fact, we show that asymptotic normality of the ISE for $\hat{f}_n$ in the signal recovery model (2) can be obtained more easily and directly via central limit theorems for quadratic forms, cf. [6], [26]. Many of our computations have to be performed in the frequency domain, which makes the analysis completely different from computations in the time domain as they occur for estimators based on kernels of finite order in nonparametric regression models on a compact interval. Those are essentially based on a Taylor expansion of the ISE and estimation of the remainder terms.

The paper is organized as follows. In Section II, we establish asymptotic normality of the ISE of $\hat{f}_n$ in case of band-limited signal $f(t)$ as well as for more general signals, which have Fourier transforms that satisfy certain tail conditions. In Section III, we develop statistical tests for checking whether a band-limited signal f in (2) belongs to a given parametric finite-dimensional submodel. These tests are based on the L_2-distance between the estimator and the parametric submodel. The asymptotic distribution of the test statistic is derived under both the null hypothesis of the validity of the parametric submodel and under fixed alternatives. This will be used to construct a test whether f follows a specific parametric form, similarly as in [7], [8], and [15]. In Section IV, we give a corresponding result in the context of regression on a compact interval. To this end, asymptotics of the ISE for the kernel regression estimator with the well-known sinc kernel for time-limited signals $f(t)$ is exploited. It is shown that the proposed method outperforms common estimators with compactly supported kernels with respect to the asymptotic relative efficiency ARE. A simulation study which investigates the finite sample behavior of the proposed tests is presented in Section V.

To conclude the introduction, let us point out that the estimator (4) can also be regarded as a spectral cutoff estimator in the direct regression model (2) (see [36]). Hence, this example serves as a prototype of the asymptotics of the ISE for spectral cutoff estimators, which naturally occur in inverse regression problems.

All proofs are deferred to an Appendix.

II. Asymptotic Normality of the ISE

Throughout this section, we assume that observations from the model (2) are available. Straightforward computation yields

$$\begin{aligned}\mathrm{ISE}(\hat{f}_n) &- \mathrm{MISE}(\hat{f}_n)\\ &= \int_{\mathbb{R}} (\hat{f}_n(t) - E\hat{f}_n(t))^2 dt - E\int_{\mathbb{R}} (\hat{f}_n(t) - E\hat{f}_n(t))^2 dt\\ &\quad + 2\int_{\mathbb{R}} (\hat{f}_n(t) - E\hat{f}_n(t))(E\hat{f}_n(t) - f(t)) dt. \qquad (5)\end{aligned}$$

First let us consider the quadratic term in (5). It turns out that it is, in fact, independent of the signal f.

Proposition 1: For any $f \in L_2(\mathbb{R})$ of finite energy in model (2) we have that

$$\int_{\mathbb{R}} (\hat{f}_n(t) - E\hat{f}_n(t))^2 dt = \frac{\tau^2\Omega}{\pi} \sum_{|j|,|k|\le n} \varepsilon_j \varepsilon_k \mathrm{sinc}(\Omega\tau(j-k)). \qquad (6)$$

If $\Omega\tau \to 0$, $\tau^3\Omega^3 n \to 0$, and $\log(n)/n = o(\Omega\tau)$ in (4), then

$$\mathrm{Var}\left[\int_{\mathbb{R}} (\hat{f}_n(t) - E\hat{f}_n(t))^2 dt\right] = \frac{4\sigma^4}{\pi}\tau^3\Omega n(1 + o(1)).$$

As for the linear term in (5), since it contains the bias $E\hat{f}_n(t) - f(t)$ as a factor, it should be asymptotically negligible as compared with the quadratic term. In order to estimate the bias, we need the following tail behavior of the signal f:

$$\text{There exist } c, r > 0 \text{ such that } |f(t)| \le c|t|^{-(r+1)}, \qquad t \in \mathbb{R}. \qquad (7)$$

This tail behavior is implied by certain smoothness assumptions on the Fourier transform F of f, see [32]. Note that, by assumption (7), $f \in L_1$. Now let us describe the asymptotic distribution of the ISE. First we consider the case of a band-limited signal.

Theorem 1: Suppose that $f \in \mathrm{BL}(\tilde{\Omega})$ satisfies (7) with some $r > 1/2$. If $\Omega \ge \tilde{\Omega}$ and if $\tau^2 n \to 0$ and $n^{2r}\tau^{2r+1} \to \infty$, then

$$\frac{1}{\sqrt{\tau^3 n}}(\mathrm{ISE}(\hat{f}_n) - \mathrm{MISE}(\hat{f}_n)) \xrightarrow{\mathcal{L}} N(0, 4\sigma^4\Omega/\pi)$$

where $N(\mu, \sigma^2)$ denotes the normal law with mean μ and variance σ^2.

Remark 1: In [32], it is shown that for $\tau = an^{-(2r+1)/(2r+2)}$, the optimal rate

$$\mathrm{MISE}(\hat{f}_n) = O(n^{-r/(r+1)}) \qquad (8)$$

is obtained. However, the assumptions of Theorem 1 are not satisfied for these τ. This phenomenon, frequently observed in nonparametric regression, is due to the fact that for the optimal rate for the MISE, the integrated bias and the integrated variance have to decay equally fast. However, for the linear term in (5) to be asymptotically negligible as compared with the quadratic term, we need that the integrated bias decays at a faster rate than the integrated variance. In our situation, choosing

$$\tau = an^{-\frac{2r+1}{2r+2}+\frac{1}{(2r+1)(2r+2)}+\varepsilon}$$

for some arbitrarily small $\varepsilon > 0$, the conditions of Theorem 1 are satisfied and from the estimates in [32], we obtain the rate

$$\mathrm{MISE}(\hat{f}_n) = O\left(n^{-\frac{r}{r+1}+\frac{1}{(r+1)(2r+1)}+2\varepsilon}\right).$$

For large r this is close to (8).

Now let us consider non-band-limited signals. Following the method proposed in [30], [32], we let $\Omega \to \infty$ for the estimator in (4), and impose additional assumptions on the tail behavior of the Fourier transform of f

$$\text{The Fourier transform } F \text{ of } f \text{ satisfies} \quad |F(\omega)| \le d|\omega|^{-(\alpha+1/2)}, \qquad |\omega| \ge 1,\ \alpha > 1. \tag{9}$$

Then we can state the following theorem.

Theorem 2: Suppose that $f \in L_2(\mathbb{R})$ satisfies (7) and (9) with $r > 1/2$, $\alpha > 1$, and $(2r-1)(2\alpha-1) > 2$. If $\tau \to 0$, $\Omega^{2\alpha}\tau \to \infty$, $\tau^2 n\Omega \to 0$, $\tau^3\Omega^3 n \to 0$, and $n^{2r}\tau^{2r+1}/\Omega \to \infty$, then

$$\frac{1}{\sqrt{\tau^3 n\Omega}}(\mathrm{ISE}(\hat{f}_n) - \mathrm{MISE}(\hat{f}_n)) \xrightarrow{\mathcal{L}} N(0, 4\sigma^4/\pi).$$

Notice that Theorem 1 could be interpreted as limit version of Theorem 2 for which $\alpha = \infty$ and Ω is constant.

III. Testing for a Parametric Form of a Band-Limited Signal

In this section, we develop a consistent test for a parametric hypothesis in form of a model U in the signal recovery model (2). Whereas extensive literature exists on parametric signal modeling (e.g., [22]), verification of the underlying parametric hypothesis has been rarely adressed in the signal processing literature, as was pointed out by a referee. To keep the presentation concise, in the following we will assume that U is a linear model, i.e., $U = \mathrm{span}\{g_1, \dots, g_m\}$ for some basis functions $g_l, l = 1, \dots, m$. Nonlinear models can be treated similarly, see [27]. As a test statistic, we will use the (squared) L_2-distance of the estimator $\hat{f}_n$ from the parametric submodel.

L_2-based methods for model tests in regression have been frequently employed in the statistics literature. In a random design regression model, the (weighted) L_2-distance of a nonparametric kernel estimator of the signal and a smoothed version of a parametric estimate was used in [15] to test the validity of a parametric model. Also, in a random design, the (weighted) L_2-norm of the signal and its derivatives was estimated in [17] by integrating the corresponding coefficients of a local polynomial estimator. In case of a regression model with fixed design on a compact interval, a test statistic can be based on the difference of a nonparametric kernel-based estimator and a parametric estimator for the variance (cf. [7], where the asymptotic distribution of the test statistic under both the hypothesis of a linear model and under fixed alternatives is derived). Typically in this context, it is assumed that the signal is sufficiently smooth, i.e., it has r continuous derivatives for some $r \ge 1$.

In this section, we obtain an analogous result in the signal recovery model (2) on the whole real line under stronger smoothness assumptions. In fact, we restrict ourselves to band-limited signals, and our arguments are based extensively on the cardinal series expansion with oversampling (3). In the general case, the error in the expansion (3) has to be estimated. For simplicity, we start by considering a simple hypothesis $H : f = f_0$. By centering the data $y'_j = y_j - f_0(\tau j)$ we may assume that $f_0 = 0$. In this case, our test statistic is

$$\hat{M}_n^2 = \int_{\mathbb{R}} (\hat{f}_n(t))^2 dt = \frac{\tau^2\Omega}{\pi} \sum_{|j|,|k| \le n} y_j y_k \mathrm{sinc}(\Omega\tau(j-k)). \tag{10}$$

Observe that $\hat{M}_n^2$ can be evaluated directly using (6) without performing a numerical integration. The next theorem gives the asymptotic behavior of $\hat{M}_n^2$.

Theorem 3: Under the hypothesis $f = 0$, if $\tau^3 n \to 0$ and $\log^2(n)/n = o(\tau)$, then

$$\frac{1}{\sqrt{\tau^3 n}}\left(\hat{M}_n^2 - \Omega\tau^2\sigma^2\pi^{-1}(2n+1)\right) \xrightarrow{\mathcal{L}} N(0, 4\sigma^4\Omega/\pi). \tag{11}$$

Under the alternative $f \ne 0$, suppose that $f \in \mathrm{BL}(\tilde{\Omega})$ satisfies (7) with $r > 1$. If $\Omega \ge \tilde{\Omega}$, $n\tau^{3/2} \to 0$, and $n^{2r}\tau^{2r+1} \to \infty$, then

$$\tau^{-1/2}\left(\hat{M}_n^2 - \|f\|^2\right) \xrightarrow{\mathcal{L}} N(0, 4\sigma^2\|f\|^2)$$

where $\|\cdot\|$ denotes the $L_2(\mathbb{R})$-norm.

Remark 2: Note that different rates appear under the hypothesis and under alternatives in Theorem 3, respectively. A similar phenomenon was observed in [7] in the context of nonparametric regression on a compact interval. In that model, the nonparametric rate nh occurs under the hypothesis, where h is a bandwidth that satisfies $h^2 n \to \infty$, and the parametric rate $n^{1/2}$ under a fixed alternative. Further, in Theorem 3 under an alternative, we get the same rate as was obtained in [33] in a central limit theorem for the *pointwise* error. Thus, in our model the $\tau^{1/2}$ rate corresponds to the parametric rate $n^{-1/2}$ in [7].

Remark 3: In general, the variance σ^2 will be unknown and thus has to be estimated (cf. Theorem 3). To this end, the following simple difference-based estimator can be used (for a detailed discussion of such estimators on a compact interval and their mean-square error (MSE)-properties cf. [9])

$$\hat{\sigma}^2 = \frac{1}{4n-2}\sum_{j=-n+1}^{n}(y_j - y_{j-1})^2. \tag{12}$$

It can be shown that in model (2)

$$E\hat{\sigma}^2 = \sigma^2 + O(\tau/n), \quad \mathrm{Var}(\hat{\sigma}^2) = O(1/n). \tag{13}$$

Observe that (13) implies that $\hat{\sigma}^2 = \sigma^2 + O_P(n^{-1/2})$. Therefore, Theorems 3 and 4 remain true if we replace σ^2 by $\hat{\sigma}^2$ in (11).

If we wish to test whether the signal f in model (2) lies in some finite-dimensional subspace U of $\mathrm{BL}(\tilde{\Omega})$, following the method proposed in [8], we can use the test statistic

$$\hat{M}_n^2 = \inf_{g \in U} \|\hat{f}_n - g\|^2.$$

Choosing any orthonormal basis $\{g_1, \ldots, g_m\}$ of U, this can be expressed as

$$\begin{aligned}\hat{M}_n^2 &= \|\hat{f}_n\|^2 - \sum_{l=1}^{m} |\langle \hat{f}_n, g_l \rangle|^2 \\ &= \frac{\tau^2 \Omega}{\pi} \sum_{|j|,|k| \le n} y_j y_k \mathrm{sinc}(\Omega\tau(j-k)) \\ &\quad - \sum_{l=1}^{m} \left(\tau \sum_{|k| \le n} y_k g_l(\tau k) \right)^2 .\end{aligned}$$

Notice that $\hat{M}_n$ can still be evaluated directly without numerical integration. Let $M^2 = \inf_{g \in U} \|f - g\|^2$.

Theorem 4: Let U be a finite-dimensional subspace of $\mathrm{BL}(\tilde{\Omega})$ such that every $g \in U$ satisfies (7) with $r > 1$. If in model (2), $f \in \mathrm{BL}(\tilde{\Omega})$ also satisfies (7) with $r > 1$ and $n\tau^{3/2} \to 0$ and $n^{2r}\tau^{2r+1} \to \infty$, then $\hat{M}_n^2$ is asymptotically unbiased for M^2 and

$$\frac{1}{\sqrt{\tau^3 n}} \left(\hat{M}_n^2 - \Omega\tau^2\sigma^2\pi^{-1}(2n+1) \right) \xrightarrow{\mathcal{L}} N(0, 4\sigma^4\Omega/\pi) \quad (14)$$

if $M^2 = 0$, and

$$\frac{1}{\sqrt{\tau}} \left(\hat{M}_n^2 - M^2 \right) \xrightarrow{\mathcal{L}} N(0, 4\sigma^2 M^2) \quad (15)$$

if $M^2 > 0$.

The proof is a rather straightforward extension of the proof of Theorem 3 and will be omitted.

Remark 4: The limit distribution (14) under the hypothesis allows now to check the model U by means of testing the hypothesis

$$H_0 : f \in U \text{ versus } K_0 : f \notin U \quad (16)$$

at a controlled error rate α before analyzing the data via the model U. To this end, $\hat{\sigma}^2$ in (12) has to be used as an estimator for σ^2 in (14) and H is rejected if

$$\frac{\frac{1}{\sqrt{\tau^3 n}} \left(\hat{M}_n^2 - \Omega\tau^2\hat{\sigma}^2\pi^{-1}(2n+1) \right)}{2\hat{\sigma}^2\sqrt{\Omega/\pi}} > U_{1-\alpha}$$

where $U_{1-\alpha}$ denotes the upper $1-\alpha$ quantile of the standard normal distribution. Note, that this yields a consistent test by (13).

Finally, the asymptotic normality in (15) can be used for two different purposes, testing hypotheses of the type

$$H_\Delta : M > \Delta \text{ versus } K_\Delta : M \le \Delta$$

and the construction of confidence intervals for M. We will not pursue this issue further and refer to [8].

IV. Tests for Time-Limited Signals

In this section, we extend our method to testing assumptions on time-limited signals, as they appear, e.g., in the detection of acoustically evoked potentials by electroencephalogram (EEG) measurements [19]. This is the classical context of nonparametric regression on a compact interval. Suppose that the signal $f(t)$ has support $\mathrm{supp}(f) \subset [-1,1]$. Assume now that noisy data of the following form are available:

$$y_k = f(k/n) + \epsilon_k, \qquad |k| \le n. \quad (17)$$

In this situation, we can use the estimator (4) with $\tau = 1/n$. Notice that except for the normalization, this estimator corresponds to a kernel regression estimator with kernel $K(x) = \sin(x)/(\pi x)$ and inverse bandwidth Ω. This kernel K is sometimes referred to as the sinc-kernel. Note that, in contrast to the setting in Section III, the signal cannot be band-limited, except if $f = 0$ (see [24], [38]). In the following, we obtain similar results for time-limited signals as in Sections II and III for the band-limited case. However, additional significant technical difficulties occur, which are due to the fact that the integrals involved are no longer taken over the whole real line. Thus, the Fourier isometry cannot be applied to integrals over sinc and indicator functions, as in Sections II and III. This complicates proofs significantly and we will only sketch the main steps in the Appendix. A comprehensive proof can be found in [3]. The next theorem gives uniform pointwise convergence of the MSE of the estimator.

Theorem 5: Suppose that in model (17), the signal f satisfies (9) with $\alpha > \frac{3}{2}$. If $\Omega = o(n^{2/3})$ for $n, \Omega \to \infty$, then uniformly on $[-1,1]$

$$E\left[(\hat{f}_{n,\Omega}(t) - f(t))^2\right] = O(\Omega^{-2\alpha+1}) + O(\Omega^3/n^2) + O(\Omega/n)$$

where

$$\hat{f}_{n,\Omega}(t) = \frac{1}{n} \sum_{|k| \le n} y_k \frac{\sin(\Omega(t-k/n))}{\pi(t-k/n)}.$$

Remark 5: Assumption (9) on the tails of the Fourier transform of F implies continuity of f on the whole real line, in particular we have $f(1) = f(-1) = 0$. This allows to show uniform convergence of our estimator on $[-1,1]$. Without such a condition kernel regression estimators without boundary correction converge to $f(x)/2$ at the boundary points, and not to the signal [10].

Remark 6: The sinc-kernel estimator achieves the same rates that can be shown to be optimal in closely related classes of functions. For example, let f be an L_1-function which satisfies (9) for $\alpha = m + 1/2$, $m \ge 2$ an integer. Then, according to Theorem 5, the pointwise MSE of the sinc-kernel estimator is $O\left(n^{-2m/(2m+1)}\right)$. The class of signals for which (9) holds with $\alpha = m + 1/2$ (see [5]) is closely related to the class $\mathcal{C}_m$ defined in [12] if some additional regularity assumptions on the mth derivative of f are made. For the class $\mathcal{C}_m$, the rate of convergence of the linear minimax risk is known to be $n^{-2m/(2m+1)}$ ([12, pp. 84–88]). Moreover, for the class of 2π-periodic functions with tail conditions on the Fourier coefficients similar to (9) it is known ([29]) that $n^{-2m/(2m+1)}$ is the minimax rate again.

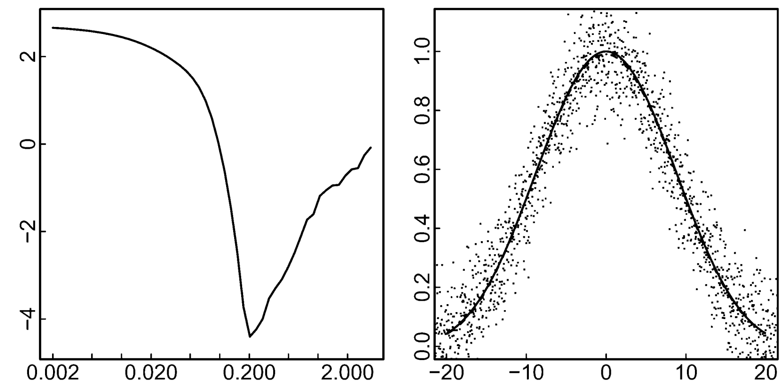

Fig. 1. Left: Logarithm of the simulated MISE between true signal and estimated signal versus τ_0. Right: Test signal (solid curve) versus estimator based on $n = 100$ (dotted curve) and $n = 1000$ (dashed curve). The dots show the set of artificial data with $n = 1000$. Note that all three curves are visually indistinguishable.

The next result describes the asymptotic distribution of the ISE for kernel regression with the $sinc$-kernel for time-limited signals, in an analogous way to Theorem 3. Consider the statistic

$$\tilde{M}_n^2 = \int_{-1}^{1} (\hat{f}_n(t))^2 dt = Y^T \tilde{A} Y$$

where $Y = (y_{-n}, \dots, y_n)^T$ and

$$\tilde{A} = (\tilde{a}_{j,k})_{|j|,|k| \le n}$$
$$\tilde{a}_{j,k} = \left(\frac{\Omega}{n\pi}\right)^2 \int_{-1}^{1} \mathrm{sinc}\left(\Omega\left(t - \frac{j}{n}\right)\right) \mathrm{sinc}\left(\Omega\left(t - \frac{k}{n}\right)\right) dt.$$

Theorem 6: Under the hypothesis $f = 0$, if $\log(n)/\sqrt{\Omega} \to 0$, and $\Omega^{3/2}/n \to 0$ as $n, \Omega \to \infty$, then

$$n\Omega^{-1/2}\left(\tilde{M}_n^2 - 2\Omega\sigma^2/(\pi n)\right) \xrightarrow{\mathcal{L}} N(0, 4\sigma^4/\pi). \tag{18}$$

Under the alternative, suppose that $f \neq 0$ satisfies (9) with $\alpha > \frac{3}{2}$. If $\ln(n)/\sqrt{\Omega} \to 0$, $\Omega^2/n \to 0$, and $\Omega^{-2\alpha}\sqrt{n} \to 0$ as $n, \Omega \to \infty$, then

$$\sqrt{n}\left(\tilde{M}_n^2 - \|f\|^2_{L^2[-1,1]}\right) \xrightarrow{\mathcal{L}} N(0, 4\sigma^2\|f\|^2_{L^2[-1,1]}).$$

Remark 7: The potential power of our test based on the statistic $\tilde{M}_n^2$ with the Fourier estimate kernel, is indicated by the consideration of local alternatives. To this end, consider the case $H : f \equiv 0$. Similarly as in [7], we obtain for the limiting variance under local alternatives of the type $f_n = (\sqrt{\Omega}/n)^{1/2} g$ the value $4\sigma^4/\pi$ as in (18). The result in [7] closely resembles (18) if the smoothing parameter of the nonparametric estimator in [7] is replaced by the multiplicative inverse of our smoothing parameter Ω. However, the regression model in [7] is $y_{j,n} = y(t_{j,n}) = m(t_{j,n}) + \varepsilon_{j,n}$, $j = 1, \dots, n$ for design points $t_{1,n}, \dots, t_{n,n} \in [0,1]$. This differs slightly from our setting, both in the number of design points (n instead of $2n+1$) and the size of the support of the design density ($[0,1]$ instead of $[-1,1]$). A close inspection of our proofs shows that if our regression model is changed into n equally spaced observations on $[0,1]$, the variance of (18) becomes $\mu_0^2 := 2\sigma^4/\pi \approx 0.64\sigma^4$.

The asymptotic variance μ_0^2 in [7, eq. (2.13)] depends on the kernel used for the nonparametric variance estimator. In the numerical simulations in [7], the Epanechnikov kernel is used. For this kernel $\mu_0^2 \approx 1.70\sigma^4$. Furthermore, for the Gauss kernel $\mu_0^2 \approx 0.81\sigma^4$, and for the sinc kernel as discussed in this paper $\mu_0^2 = 2\sigma^4/\pi$, thus, the variance for our test based on the sinc kernel is formally recovered by [7, eq. (2.13)]. However, note that for the Gauss kernel and the sinc kernel the assumption of a compactly supported kernel does not hold, so the results in [7] cannot be applied to these kernels. Hence, our result extends the theory by sampling-based methods to the sinc kernel which outperforms tests based on the kernels mentioned above. In particular, the asymptotic relative efficiency of the test based on the sinc kernel is ≈ 2.67 as compared to the test based on the Epanechnikov kernel, and ≈ 1.27 if the Gauss kernel is used. Note, that asymptotically this corresponds to the ratio of sample sizes required to achieve the same power, i.e., use of the sinc kernel reduces the required sample size compared to the Epanechnikov kernel by a relative amount of ≈ 2.67 and to the Gauss kernel by ≈ 1.27, respectively.

V. Simulation Results

In this section, we investigate the finite-sample behavior of the tests presented in Section III, which are based on asymptotic theory. In Section V-A, we comment on the selection of the parameters Ω, τ which occur in the estimator $\hat{f}_n$. Furthermore, in Section V-B, we present simulations of the distribution of $\hat{M}_n^2$ for finite sample size, both under the hypothesis $f = 0$ and under the alternative of a particular nonzero band-limited signal.

A. Choosing the Parameters

In order to compute the estimator $\hat{f}_n$, the parameters τ and Ω have to be chosen. These need to be fixed prior to application of the estimator to a given set of observations. In this subsection, we consider noisy data of the form (2), where the signal f is the band-limited function

$$f_1(t) = (\mathrm{sinc}\Omega_f t)^4 \in \mathrm{BL}(4\Omega_f), \qquad \text{where } \Omega_f = 0$$

and determine suitable values for τ and Ω. As sample size we consider $N_{\mathrm{sample}} = 2n + 1 = 201$ and 2001, and the errors ε_k are taken as i.i.d. normally distributed with zero mean and variance $\sigma^2 = 0.01$.

First, we chose $\Omega = 0.4$, which is the smallest value such that $f_1 \in \mathrm{BL}(\Omega)$. Now let us consider how to choose τ, which depends on n. The left plot in Fig. 1 presents the simulated

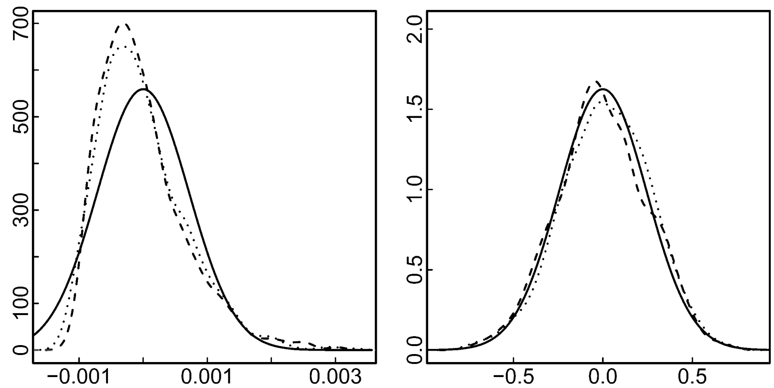

Fig. 2. (Theoretical) asymptotic normal density of $\hat{M}_n^2$ (solid curve) versus simulated density for sample size $n = 100$ (dashed curve) and $n = 1000$ (dotted curve). The left plot shows the distribution under the hypothesis $f = 0$, the right plot under the alternative if the test signal f_1 is present in the data.

MISE for estimation of $f_1(t)$ from 20 sets of artificial data with $n = 100$ for a range of different values of τ. We choose $\tau_0 = 0.2$, subsequently writing τ_0 for the value of τ for samples with $n = 100$. The very good quality of recovery of the signal f_1 in these simulations is shown in the right plot of Fig. 1, where typical estimates $\hat{f}$ based on $\tau_0 = 0.1$ for $n = 100$ and $n = 1000$ are shown. Since τ depends on n, we scale τ_0 for simulations with $n \neq 100$ as $\tau(n) = \tau_0 \cdot (100/n)^{4/5}$, in accordance with the conditions in Theorems 2 and 3.

B. Finite Sample Behavior of $\hat{M}_n^2$

In this subsection, simulations of the distribution of $\hat{M}_n^2$ for finite sample size are reported. First, we consider pure noise, i.e., generated from the signal $f = 0$. In the left plot in Fig. 2, the theoretical asymptotic normal distribution together with simulated finite sample distributions of $\hat{M}_n^2$ for $n = 100$ (dashed curve) $n = 1000$ (dotted curve) is displayed. The approximation of the asymptotic normal distribution is not too satisfactory even for the already rather large sample size ($n = 1000$). This parallels findings for related test statistics (see [8], [27]). Here, bootstrap approximations or second-order corrections can be used to improve (see, e.g., [27]). Under the alternative f_1, as shown in the right plot in Fig. 2, the approximation by the theoretical asymptotic distribution is rather accurate already at a moderate sample size ($n = 100$).

VI. Concluding Remarks and Possible Extensions

In this paper, tests for parametric assumptions on band- and time-limited signals which are observed under noise have been constructed. As a test statistic, the L_2-distance of an estimator based on the WS sampling theorem with oversampling to the parametric model is used. The asymptotic distribution of the test statistic is derived both under the hypothesis of the validity of the parametric model and under fixed alternatives. This allows in particular to test whether the signal f is close in the L_2 distance to the parametric model, at a controlled error rate. The asymptotics are valid under certain rates $\tau \to 0$ and $\Omega \to \infty$, however, it is not immediately clear how to choose the parameters for some fixed sample size. In Section V some suggestions are given, but it would be interesting to investigate fully data-driven methods like cross-validation. As pointed out by a reviewer, it might also be interesting to investigate whether the theory could be established for a fixed sampling rate τ. Simulations of the finite sample behavior of the test indicate that while under the alternative, the approximation by the normal limit law is rather accurate already for moderate sample sizes, this is not the case under the hypothesis. It would be of interest to improve the approximation under the hypothesis by some resampling procedure.

There are several possible extensions of the methodology proposed in this paper. An important one is the comparison of signals under various input conditions. In this case, independently d different samples

$$y_{kj} = f_j(\tau k) + \varepsilon_{kj}, \qquad |k| \leq n,\ j = 1, \ldots, d$$

are observed and it is to be tested whether $f_1 = \cdots = f_d$. This can be based on the pairwise comparison between all samples (cf. [28]). For $d = 2$, e.g., the observations $Y_{1j} - Y_{2j}$ will simply be used in the statistic $\hat{M}_n^2$ in (10).

For further investigations it might be of interest to weaken the assumptions on the tail behavior of the signal. Based on this, tests for signals which are neither band- nor time-limited could be constructed in a similar fashion. As an example, consider the problem to decide whether an exponentially damped sinusoidal model holds ([1], [22], or [37]), where

$$f(k) = \sum_{l=1}^{m} \alpha_l e^{s_l k}. \tag{19}$$

Here α_l and s_l are unknown (complex) numbers, such that $\mathrm{Re}(s_k) < 0$. Note that this implies integrability of the signals $f(\cdot)$. Further, m is assumed to be fixed. In our terminology this would be a parametric model with parameters α_l, s_l; the α_l are linear parameters, the s_l nonlinear.

It would also be of interest to consider a more general dependent noise process. Finally, let us stress that all signals considered in this paper are of finite energy (i.e., in $L_2(\mathbb{R})$). However, several frequently encountered signals like cosine functions do not satisfy this requirement, and a theory that covers such signals would be of much practical interest, as well.

Appendix

Recall that the Fourier transform of a signal $f \in L_2(\mathbb{R})$ is given by

$$F = \mathcal{F}(f)(\omega) = \int_{\mathbb{R}} f(t)e^{-it\omega}dt$$

so that the inverse transform is given by

$$\mathcal{F}^{-1}(F)(t) = \frac{1}{2\pi}\int_{\mathbb{R}} F(\omega)e^{it\omega}d\omega.$$

Hence, the Fourier transform of the estimator $\hat{f}_n(t)$ in (4) is given by

$$\hat{F}_n(\omega) = \tau \sum_{|k|\le n} y_k 1_{[-\Omega,\Omega]}(\omega)e^{-i\omega k\tau}. \tag{20}$$

A. *Proof of Proposition 1*

From Parseval's equation and (20)

$$\begin{aligned}\int_{\mathbb{R}}(\hat{f}_n(t) - E\hat{f}_n(t))^2dt &= \frac{1}{2\pi}\int_{\mathbb{R}}|\hat{F}_n(\omega) - E\hat{F}_n(\omega)|^2d\omega\\ &= \frac{1}{2\pi}Z_n^TAZ_n \end{aligned}\tag{21}$$

where $Z_n = (\epsilon_{-n},\dots,\epsilon_n)^T$ and $A = (a_{j,k})_{|j|,|k|\le n}$, $a_{j,k} = 2\tau^2\Omega\,\mathrm{sinc}\,(\Omega\tau(j-k))$, $|j|,|k|\le n$, which proves (6). The expectation of the quadratic form in (21) is given by

$$EZ_n^TAZ_n = \tau^2\sigma^2 2\Omega(2n+1) \tag{22}$$

and, thus, we can write

$$\begin{aligned}&Z_n^TAZ_n - EZ_n^TAZ_n\\ &\quad= \sum_{\substack{|j|,|k|\le n\\ j\ne k}} a_{j,k}\epsilon_j\epsilon_k + 2\Omega\tau^2\sum_{|j|\le n}\left(\epsilon_j^2-\sigma^2\right) = T_1 + T_2.\end{aligned}$$

Evidently

$$ET_1 = ET_2 = 0, \quad \mathrm{Cov}(T_1,T_2) = ET_1T_2 = 0$$

and

$$ET_2^2 = 4\Omega^2\tau^4(2n+1)E\left(\epsilon_1^2-\sigma^2\right)^2. \tag{23}$$

Moreover

$$\begin{aligned}ET_1^2 &= 2\sigma^4\sum_{|j|,|k|\le n, j\ne k} a_{j,k}^2\\ &= 16\sigma^4\tau^4\Omega^2\Bigg((2n+1)\sum_{j=1}^{2n}\mathrm{sinc}^2(\Omega\tau j)\\ &\qquad -\sum_{j=1}^{2n} j\mathrm{sinc}^2(\Omega\tau j)\Bigg).\end{aligned}\tag{24}$$

In order to compute the asymptotic variance of T_1, in a first step, we replace the sums in (24) by integrals and in a second step we estimate the approximation error. For the first sum this gives

$$\int_{1/2}^{2n+1/2}\mathrm{sinc}^2(\Omega\tau t)dt = \frac{1}{\Omega\tau}\int_{\Omega\tau/2}^{(2n+1/2)\Omega\tau}\mathrm{sinc}^2(u)du. \tag{25}$$

Since $\Omega\tau\to 0$ and $n\Omega\tau\to\infty$

$$\int_{\Omega\tau/2}^{(2n+1/2)\Omega\tau}\mathrm{sinc}^2(u)du \to \int_0^\infty \mathrm{sinc}^2(u)du = \pi/2. \tag{26}$$

For the second sum we obtain

$$\begin{aligned}\int_{1/2}^{2n+1/2} t\mathrm{sinc}^2(\Omega\tau t)dt &= \frac{1}{\Omega^2\tau^2}\int_{\Omega\tau/2}^{(2n+1/2)\Omega\tau} u\mathrm{sinc}^2(u)du\\ &= O(\log(\Omega\tau n)/(\Omega^2\tau^2)).\end{aligned}\tag{27}$$

The approximation errors are estimated in Lemma 1. Collecting terms from (23)–(29) gives

$$\begin{aligned}Var\left(Z_n^TAZ_n\right) &= 16\sigma^4\tau^3\Omega(2n+1)(\pi/2 + o(1)\\ &\quad + O(\log(\Omega\tau n)/(\Omega\tau n))\\ &\quad + O((n\tau^3\Omega^3)^{1/2}) + O(\log n/(\Omega\tau n))).\end{aligned}$$

Taking into account the factor in (21) yields the proposition. □

The following lemma provides the missing estimates of the approximation errors used in the above proof.

Lemma 1: We have

$$\left|\sum_{k=1}^{2n}\mathrm{sinc}^2(\Omega\tau k) - \int_{1/2}^{2n+1/2}\mathrm{sinc}^2(\Omega\tau t)dt\right| = O((n\tau\Omega)^{1/2}) \tag{28}$$

and

$$\begin{aligned}&\left|\sum_{k=1}^{2n} k\mathrm{sinc}^2(\Omega\tau k) - \int_{1/2}^{2n+1/2} t\mathrm{sinc}^2(\Omega\tau t)dt\right|\\ &\qquad = O(\log(n)/(\Omega^2\tau^2)).\end{aligned}\tag{29}$$

Proof: For $\eta > 0$, we apply Lemma 4 in [33] to the function $f(t) = \mathrm{sinc}(\eta(t+n)) \in BL(\eta)$ with $\tau = 1$ and obtain

$$\begin{aligned}\sum_{k=1}^{2n}\int_{k-1/2}^{k+1/2}(\mathrm{sinc}(\eta t) - \mathrm{sinc}(\eta k))^2dt &\le \eta^2\int_{\mathbb{R}}\mathrm{sinc}^2(\eta t)dt/\pi^2\\ &= \eta/\pi.\end{aligned}\tag{30}$$

Thus,

$$\begin{aligned}&\left|\sum_{k=1}^{2n}\mathrm{sinc}^2(\Omega\tau k) - \int_{1/2}^{2n+1/2}\mathrm{sinc}^2(\Omega\tau t)dt\right|\\ &\quad\le \sum_{k=1}^{2n}\int_{k-1/2}^{k+1/2} r(\Omega,\tau,t,k)|\mathrm{sinc}(\Omega\tau t) + \mathrm{sinc}(\Omega\tau k)|dt\\ &\quad\le 2\sup_{x\in\mathbb{R}}|\mathrm{sinc}(x)|\sqrt{2n}\left(\sum_{k=1}^{2n}\int_{k-1/2}^{k+1/2} r(\Omega,\tau,t,k)^2dt\right)^{1/2}\\ &\quad= O((n\tau\Omega)^{1/2})\end{aligned}$$

where $r(\Omega,\tau,t,k) = |\mathrm{sinc}(\Omega\tau t) - \mathrm{sinc}(\Omega\tau k)|$ and we used (30) in the last step with $\eta = \Omega\tau$. Moreover, we get the expressions at the bottom of the page. In the next lemma, we estimate the linear part in (5). □

Lemma 2: In the band-limited case, suppose that $f \in \mathrm{BL}(\tilde{\Omega})$ satisfies (7) for some $r > 1/2$. If $\Omega \geq \tilde{\Omega}$ in the estimator (4) and if $\tau^2 n \to 0$ and $n^{2r}\tau^{2r+1} \to \infty$, then

$$\int_{\mathbb{R}} (\hat{f}_n(t) - E\hat{f}_n(t))(E\hat{f}_n(t) - f(t))dt = o_P((\tau^3 n)^{1/2}).$$

In the non-band-limited case, suppose that $f \in L_2(\mathbb{R})$ satisfies (7) and (9) with $r > 1/2$ and $\alpha > 1$. If $\tau \to 0$, $\Omega^{2\alpha}\tau \to \infty$, $\tau^2 n\Omega \to 0$, $\tau^4\Omega^5 n \to 0$, and $n^{2r}\tau^{2r+1}/\Omega \to \infty$, then

$$\int_{\mathbb{R}} (\hat{f}_n(t) - E\hat{f}_n(t))(E\hat{f}_n(t) - f(t))dt = o_P((\tau^3 \Omega n)^{1/2}).$$

Proof: From the Cauchy–Schwarz inequality

$$\left|\int_{\mathbb{R}} (\hat{f}_n(t) - E\hat{f}_n(t))(E\hat{f}_n(t) - f(t))dt\right| \leq (V(\hat{f}_n))^{1/2}\left(\mathrm{IBIAS}^2(\hat{f}_n)\right)^{1/2}$$

where

$$V(\hat{f}_n) = \int_{\mathbb{R}} \left(\hat{f}_n(t) - E\hat{f}_n(t)\right)^2 dt$$

and

$$\mathrm{IBIAS}^2(\hat{f}_n) = \int_{\mathbb{R}} \left(E\hat{f}_n(t) - f(t)\right)^2 dt.$$

From (22)

$$V(\hat{f}_n) = O_P(\tau^2 n\Omega).$$

Furthermore, from the estimates of the integrated bias in [32] (Theorem 2 for the band-limited and Theorem 3 for the non-band-limited case) we get

$$\mathrm{IBIAS}^2(\hat{f}_n) = o(\tau). \qquad \square$$

B. Proofs of Theorems 1 and 2

From Lemma 2, it follows that the linear part in (5) is $o_P\left((\tau^3\Omega n)^{1/2}\right)$. Furthermore, from (23), the diagonal part T_2 of the quadratic form in (23) is

$$O_P\left((\tau^4\Omega^2 n)^{1/2}\right) = o_P\left((\tau^3\Omega n)^{1/2}\right)$$

as well. Moreover, in both cases from the assumptions, it follows that $\log^2(n)/n = o(\tau)$, and Proposition 1 applies. Thus, it remains to prove asymptotic normality of T_1. To this end we apply [6, Theorem 5.2]. By a straightforward calculation

$$\frac{1}{\tau^3 n\Omega} \max_{|j|\leq n} \sum_{|k|\leq n, k\neq j} a_{j,k}^2 = O(1/n) \tag{31}$$

therefore, Assumptions 1) and 2) of [6, Theorem 5.2] are satisfied with $K(n) = \tau^{-1/4}$. Next we use the fact that the spectral radius $\rho(A)$ of a symmetric matrix A is bounded from above by any matrix operator norm. Therefore,

$$\begin{aligned} \frac{1}{\sqrt{\tau^3 n\Omega}}\rho(A) &\leq \frac{1}{\sqrt{\tau^3 n\Omega}} \max_{|j|\leq n} \sum_{|k|\leq n, k\neq j} |a_{j,k}| \\ &= O(\log(n)/(n\tau)^{1/2}) = o(1) \end{aligned} \tag{32}$$

which yields Assumption 3) in [6]. This concludes the proof of Theorems 1 and 2. □

C. Proof of Theorem 3

Notice that for $f = 0$, the estimator $\hat{f}_n$ is unbiased. The assumptions of Proposition 1 are satisfied, moreover, both terms in (31) and (32) tend to zero, and [6, Theorem 5.2] applies again. Now let us consider the case of $f \neq 0$. Similarly as in the proof of Proposition 1

$$\int_{\mathbb{R}} (\hat{f}_n(t))^2 dt = \frac{1}{2\pi} Y_n^T A Y_n \tag{33}$$

where A is as before and $Y_n = (y_{-n}, \ldots, y_n)^T$. We have

$$\begin{aligned} EY_n^T A Y_n &= 2\Omega\tau^2\sigma^2(2n+1) \\ &\quad + 2\Omega\tau^2 \sum_{|j|,|k|\leq n} f(\tau j) f(\tau k)\mathrm{sinc}(\Omega\tau(j-k)). \end{aligned} \tag{34}$$

$$\begin{aligned} &\left|\sum_{k=1}^{2n} k\,\mathrm{sinc}^2(\Omega\tau k) - \int_{1/2}^{2n+1/2} t\,\mathrm{sinc}^2(\Omega\tau t)dt\right| \\ &\leq \sum_{k=1}^{2n} \int_{k-1/2}^{k+1/2} |t\,\mathrm{sinc}^2(\Omega\tau t) - k\,\mathrm{sinc}^2(\Omega\tau k)|dt \\ &\leq \frac{1}{\Omega^2\tau^2} \sum_{k=1}^{2n} \int_{k-1/2}^{k+1/2} \left[\left(\left|\frac{\sin^2(\Omega\tau t)}{t} - \frac{\sin^2(\Omega\tau t)}{k}\right|\right) + \left(\left|\frac{\sin^2(\Omega\tau t)}{k} - \frac{\sin^2(\Omega\tau k)}{k}\right|\right)\right] dt \\ &= \frac{1}{\Omega^2\tau^2}(O(1) + O(\log(n))). \end{aligned}$$

From the sampling theorem (3) and the tail behavior of f, we get uniformly in $|j| \le n$

$$\begin{aligned}\left|\tau \sum_{|k|\le n} f(\tau k)\frac{\sin(\Omega\tau(j-k))}{\tau(j-k)} - \pi f(\tau j)\right| &= \left|\tau \sum_{|k|>n} f(\tau k)\frac{\sin(\Omega\tau(j-k))}{\tau(j-k)}\right| \\ &\le \Omega\tau \sum_{|k|>n} |f(\tau k)| = O((n\tau)^{-r}). \end{aligned} \quad (35)$$

Therefore,

$$\begin{aligned} 2\tau^2 \sum_{|j|,|k|\le n} f(\tau j) f(\tau k)\frac{\sin(\Omega\tau(j-k))}{\tau(j-k)} &= 2\tau \sum_{|j|\le n} f(\tau j)(\pi f(\tau j) + O((n\tau)^{-r})) \\ &= 2\tau\pi \sum_{|j|\le n} f(\tau j)^2 + O((n\tau)^{-r}) \\ &= 2\pi\|f\|^2 + O(\sqrt{\tau^3 n}) + O((n\tau)^{-r}) \end{aligned} \quad (36)$$

where we used [32, Lemma 4] in the last step. Next, we decompose the quadratic form into

$$\begin{aligned} Y_n^T A Y_n - E Y_n^T A Y_n &= \sum_{|j|,|k|\le n, j\ne k} a_{j,k}(y_j y_k - f(\tau j) f(\tau k)) \\ &\quad + 2\Omega\tau^2 \sum_{|j|\le n} (y_j^2 - \sigma^2 - f(\tau j)^2) \\ &= T_1 + T_2. \end{aligned}$$

From (23)

$$T_2 = O_P(\tau^2\sqrt{n}).$$

Setting $f_n = (f(-\tau n), \dots, f(\tau n))^T$, $Z_n = Y_n - f_n$, and $B = (b_{j,k})_{|j|,|k|\le n}$ with $b_{j,k} = a_{j,k}(1-\delta_{j,k})$, where $\delta_{j,k}$ denotes the Kronecker symbol, we decompose T_1 as follows:

$$T_1 = Z_n^T B Z_n + 2 Z_n^T B f_n = T_{1,1} + 2T_{1,2}. \quad (37)$$

Evidently

$$E T_{1,1} = E T_{1,2} = \mathrm{Cov}(T_{1,1}, T_{1,2}) = 0.$$

From Proposition 1, $T_{1,1} = O_P\big((\tau^3 n)^{1/2}\big)$. We have

$$T_{1,2} = \sum_{|j|\le n} \epsilon_j \sum_{|k|\le n, j\ne k} a_{j,k} f(\tau k).$$

Using (35) once again we get

$$\begin{aligned} Var(T_{1,2}) &= \sum_{|j|\le n} \sigma^2 \left(\sum_{|k|\le n, j\ne k} a_{j,k} f(\tau k)\right)^2 \\ &= 4\sigma^2\tau^2 \sum_{|j|\le n} (f(\tau j)\pi + O((n\tau)^{-r}))^2 \\ &= 4\sigma^2\pi^2\tau\|f\|^2 + o(\tau) \\ &\quad + O(n^{-r}\tau^{1-r}) + O(n^{1-2r}\tau^{2-2r}) \\ &= 4\sigma^2\pi^2\tau\|f\|^2(1+o(1)). \end{aligned}$$

From (33), (34), (36), and the above estimates on T_2 and $T_{1,1}$

$$\begin{aligned} \hat{M}_n^2 &= O(\tau^2 n) + \|f\|^2 + o(\tau) + O((n\tau)^{-r}) \\ &\quad + O_P(\tau^2\sqrt{n}) + O_P(\sqrt{\tau^3 n}) + \pi^{-1} T_{1,2} \end{aligned}$$

therefore it will suffice to show asymptotic normality of $T_{1,2}/\sqrt{\tau}$. To this end, we apply Lyapounov's theorem. From (35) together with a straightforward calculation

$$\begin{aligned} \frac{1}{Var(T_{1,2})^2} \sum_{|j|\le n} E(\epsilon_j)^4 \left(\sum_{|k|\le n, j\ne k} a_{j,k} f(\tau k)\right)^4 &\le \frac{C}{\tau^2} \sum_{|j|\le n} \tau^4 (f(\tau j)\pi + O((n\tau)^{-r}))^4 \to 0. \quad \square \end{aligned}$$

Remark 8: Note that under the hypothesis, the term $T_{1,2}$ in (30) vanishes. Thus, the quadratic term $T_{1,1}$ determines the asymptotics and a result like [6, Theorem 5.2] for random quadratic forms has to be applied. In contrast, under the alternative, the linear term $T_{1,2}$ dominates the asymptotics. Therefore, it is no longer possible to use the above mentioned result, instead one simply may apply Lyapunov's central limit theorem.

D. Proof of Theorem 5

The proof follows mostly along the lines of the [33, proof of Theorem 1], which deals with the band-limited case. Their estimates are based on a lemma (Lemma 4) which only applies to band-limited signals, and which therefore cannot be used in our setting. Instead, we invoke [32, Lemma 3]. However, the details are rather cumbersome and can be found in [3]. $\square$

E. Proof of Theorem 6

The expectation of $\tilde{M}_n^2$ is given by

$$E\tilde{M}_n^2 = \sum_{|j|,|k|\le n} \tilde{a}_{j,k} f(t_j) f(t_k) + \sigma^2 \mathrm{tr}(\tilde{A}).$$

Tedious but straightforward computations yield that

$$\mathrm{tr}(\tilde{A}) = 2\Omega/(\pi n) + o(\sqrt{\Omega}/n)$$

and that

$$\sum_{|j|,|k|\le n} \tilde{a}_{j,k} f(t_j) f(t_k) = \|f\|^2_{L^2[-1,1]} + o(n^{-1/2}).$$

The variance of $\tilde{M}_n^2$ is computed as in the proofs of Proposition 1 and Theorem 3. If $f = 0$, the dominating term in the variance is $2\sigma^4 \sum_{|j|,|k|\le n, j\ne k} a_{j,k}^2$, otherwise it is $4\sigma^2(\tilde{A} f_n)^T(\tilde{A} f_n)$, where

$$f_n = (f(-1), f((-n+1)/n), \dots, f(1))^T.$$

Technical difficulties arise since the entries of $\tilde{A}$ can no longer be calculated explicitely by Fourier transformation, because we integrate over a finite interval. Therefore, we have to determine the asymptotic behavior of sums $\Sigma_{|k|,|l|\leq n}$ over the squares of integrals of type

$$\begin{aligned}\int_{-1}^{1} \operatorname{sinc}\left(\Omega\left(t-\frac{k}{n}\right)\right) \operatorname{sinc}\left(\Omega\left(t-\frac{l}{n}\right)\right) dt \\ = \frac{1}{\Omega}\int_{-\Omega}^{\Omega} \operatorname{sinc}\left(t-\frac{\Omega k}{n}\right) \operatorname{sinc}\left(t-\frac{\Omega l}{n}\right) dt \\ = \frac{1}{\Omega}\left(\int_{-\infty}^{\infty} \operatorname{sinc}\left(t-\frac{\Omega k}{n}\right) \operatorname{sinc}\left(t-\frac{\Omega k}{n}\right) dt \right. \\ \left. - \int_{|t|\geq\Omega} \operatorname{sinc}\left(t-\frac{\Omega k}{n}\right) \operatorname{sinc}\left(t-\frac{\Omega l}{n}\right) dt\right) \\ = \frac{1}{\Omega}\left(\operatorname{sinc}\left(\frac{\Omega}{n}(k+l)\right) + c_{ij}\right).\end{aligned}$$

To this end, we show that

$$\sum_{|k|,|l|\leq n} \operatorname{sinc}^2\left(\frac{\Omega}{n}(k+l)\right) = \frac{4n^2}{\Omega} + o\left(\frac{n^2}{\Omega}\right)$$

and hence,

$$\sum_{|k|,|l|\leq n} c_{ij}^2 = O\left(n^2/\Omega\right).$$

Thus, the tails are negligible. Finally, asymptotic normality under the hypothesis follows again from [6, Theorem 5.2], while under the alternative the Lyapounov central limit theorem is applied. □

ACKNOWLEDGMENT

The authors would like to thank M. Pawlak and U. Stadtmüller for pointing out some helpful references and three referees for their valuable comments.

REFERENCES

[1] G. K. Allu, “Estimating the Parameters of Exponentially Damped Sinusoids in Noise,” University of Rhode Island, Kingston, RU, Tech. Rep., 2003.

[2] P. J. Bickel and M. Rosenblatt, “On some global measures of the deviations of density function estimate,” *Ann. Statist.*, vol. 1, pp. 1071–1095, 1973.

[3] N. Bissantz and H. Holzmann, “On weak convergence of the squared L_2-norm of the sinc-kernel estimator for regression on a compact interval,” J. Statist. Plann. Inf., to be published.

[4] P. L. Butzer and R. L. Stens, “Sampling theory for not necessarily band-limited functions: A historical overview,” *SIAM Rev.*, vol. 32, pp. 40–53, 1992.

[5] K. Chandrasekharan, *Classical Fourier Transforms*. Berlin, Germany: Springer-Verlag, 1989.

[6] P. de Jong, “A central limit theorem for generalized quadratic forms,” *Probab. Theory Related Fields*, vol. 75, pp. 261–277, 1987.

[7] H. Dette, “A consistent test for the functional form of a regression based on a difference of variance estimators,” *Ann. Statist.*, vol. 27, pp. 1012–1040, 1999.

[8] H. Dette and A. Munk, “Validation of linear regression models,” *Ann. Statist.*, vol. 26, pp. 778–800, 1998.

[9] H. Dette, A. Munk, and T. Wagner, “Estimating the variance in nonparametric regression—What is a reasonable choice?,” *J. Roy. Statist. Soc. B*, vol. 60, pp. 751–764, 1998.

[10] S. Efromovich, *Nonparametric Curve Estimation*. New York: Springer-Verlag, 1999.

[11] R. L. Eubank, *Nonparametric Regression and Spline Smoothing*, 2nd ed. New York: Marcel Dekker, 1999.

[12] J. Fan and I. Gijbels, *Local Polynomial Modeling and its Applications*. London, U.K.: Chapman & Hall, 1996.

[13] P. Hall, “Central limit theorem for integrated square error of multivariate nonparametric density estimators,” *J. Multivariate Anal.*, vol. 14, pp. 1–16, 1984.

[14] ——, “Integrated square error properties of kernel estimators of regression functions,” *Ann. Statist.*, vol. 12, pp. 241–260, 1984.

[15] W. Härdle and E. Mammen, “Comparing nonparametric versus parametric regression fits,” *Ann. Statist.*, vol. 21, pp. 1926–1947, 1993.

[16] J. R. Higgins, “Five short stories about the cardinal series,” *Bull. Amer. Math. Soc.*, vol. 12, pp. 45–89, 1985.

[17] L. S. Huang and J. Fan, “Nonparametric estimation of quadratic regression functionals,” *Bernoulli*, vol. 5, pp. 927–949, 1999.

[18] D. A. Ioannides, “Integrated square error of nonparametric estimators of regression function: The fixed design case,” *Statist. Probab. Lett.*, vol. 15, pp. 85–94, 1992.

[19] J. T. Jacobson, *Principles and Applications in Auditory Evoked Potentials*. Boston, MA: Allyn and Bacon, 1994.

[20] A. J. Jerri, “The Shannon sampling theorem—Its various extensions and applications: A tutorial review,” *Proc. IEEE*, vol. 65, pp. 1565–1596, 1977.

[21] T. Kailath and H. V. Poor, “Detection of stochastic processes,” *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2230–2259, Oct. 1998.

[22] S. M. Kay, *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[23] T. K. Kim, “Central limit theorem for quadratic errors of nonparametric estimators,” *J. Statist. Planning Inference*, vol. 64, pp. 193–204, 1997.

[24] R. J. Marks II, *Introduction to Shannon Sampling and Interpolation Theory*. New York: Springer-Verlag, 1991.

[25] R. J. Marks II, Ed., *Advanced Topics in Shannon Sampling and Interpolation Theory*. New York: Springer-Verlag, 1993.

[26] T. Mikosch, “Functional limit theorems for random quadratic forms,” *Stochastic Process. Appl.*, vol. 37, pp. 81–98, 1991.

[27] A. Munk, “Testing the goodness of fit of parametric regression models with random Toeplitz forms,” *Scand. J. Statist.*, vol. 29, pp. 501–533, 2002.

[28] A. Munk and H. Dette, “Nonparametric comparison of regression functions: Exact and asymptotic theory,” *Ann. Statist.*, vol. 26, pp. 2339–2368, 1998.

[29] A. Munk and F. Ruymgaart, “Minimax rates for estimating the variance and its derivatives in nonparametric regression,” *Aust. N. Z. J. Statist.*, vol. 44, pp. 479–488, 2002.

[30] M. Pawlak, E. Rafajłowicz, and A. Krzyżak, “Postfiltering versus prefiltering for signal recovery from noisy samples,” *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3195–3212, Dec. 2003.

[31] M. Pawlak and U. Stadtmüller, “Recovering band-limited signals under noise,” *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1425–1438, Sep. 1996.

[32] ——, “Kernel regression estimators for signal recovery,” *Statist. Probab. Lett.*, vol. 31, pp. 185–198, 1997.

[33] ——, “Nonparametric estimation of a class of smooth functions,” *J. Nonparametr. Statist.*, vol. 8, pp. 149–183, 1997.

[34] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.

[35] T. S. Rappaport, *Wireless Communications: Principles & Practice*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 2002.

[36] A. C. M. van Rooij, F. H. Ruymgaart, and W. R. van Zwet, “Asymptotic efficiency of inverse estimators,” *Theory Probab. Appl.*, vol. 44, pp. 826–844, 1999.

[37] P. Stoica, H. Li, and J. Li, “Amplitude estimation of sinusoidals signals: Survey, new results, and an application,” *IEEE Trans. Signal Process.*, vol. 48, no. 2, pp. 338–352, Feb. 2000.

[38] A. I. Zayed, *Advances in Shannon's Sampling Theory*. Boca Raton, FL: CRC, 1993.