Statistical inference for mean and variance of oscillatory processes

Dylan Glotzer, Vladas Pipiras*
University of North Carolina

Vadim Belenky, Kenneth M. Weems†
David Taylor Model Basin

June 6, 2018

Abstract

When constructing confidence intervals for the mean and variance of a stationary continuous-time stochastic process, two approaches have been considered in the literature: one based on the so-called long-run variance of the process and its square, and the other based on the so-called self-normalization. These approaches are revisited here in the context of random oscillatory processes such as random (non-)linear oscillators and related models with particular attention to the problem of estimating the non-zero long-run variances of the processes. The case of the zero long-run variance, which has been studied and is quite different, is also considered. The approaches are extended to the situation where multiple independent records of the stochastic process are available, for example, by introducing an estimator of the long-run variance which improves on other natural candidate estimators. Finally, a simulation study is provided to assess the performance of the proposed methods in estimating the long-run variances and constructing confidence intervals, and a data example is considered.

AMS subject classifications. Primary: 62M10, 62M15, 62G05. Secondary: 60H10, 60K40, 70L05.

Keywords and phrases: (non-)linear random oscillator, long-run variance, self-normalization, degenerate case, confidence interval, kernel function, bandwidth, multiple records.

1 Introduction

Let $\{X_t\}_{t\in\mathbb{R}}$ be a stationary continuous-time stochastic process with mean $\mu(X) = \mathbb{E}X_t$ and variance $\sigma^2(X) = \mathrm{Var}(X_t)$. We are interested here in the processes representing random oscillatory systems, with a number of examples (linear random oscillators, etc.) discussed in Appendix A. The motivating real applications of our particular interest include processes related to ship motions (e.g. ship roll, pitch and yaw dynamics, bending moments, etc.), but potential applications would naturally arise in a range of other areas where random oscillatory systems play a fundamental role (see e.g. Gitterman (2005), Neimark and Landa (2012), Liang and Lee (2015)). The basic problem addressed in this work is providing confidence intervals for the mean $\mu(X)$ and the variance $\sigma^2(X)$

*Dept. of Statistics and Operations Research, UNC at Chapel Hill, CB#3260, Hanes Hall, Chapel Hill, NC 27599, USA, *dglotzer@email.unc.edu, pipiras@email.unc.edu*.

†Naval Surface Warfare Center Carderock Division (NSWCCD), 9500 MacArthur Blvd, W. Bethesda, MD 20817, USA, *vadim.belenky@navy.mil*, *kenneth.weems@navy.mil*.

(and consequently for the standard deviation $\sigma(X)$ as well) from the observed continuous-time sample X_t, $t \in [0,T]$, or the corresponding discrete sample $X_{\Delta k}$, $k = 1, \ldots, T/\Delta$. Though it should be kept in mind that the approach discussed here also applies to other quantities of possible interest, e.g. autocovariance at a given lag (see Remarks 3.1 and 3.2 below).

Focusing on the continuous-time sample X_t, $t \in [0,T]$, and on estimating the mean $\mu(X)$, consider the sample mean

$$\overline{X}_T = \frac{1}{T}\int_0^T X_s ds. \tag{1.1}$$

Its normalized asymptotic variance is

$$\lim_{T\to\infty} T \operatorname{Var}(\overline{X}_T) = \int_{-\infty}^{\infty} \Gamma_X(u) du =: \Pi(X), \tag{1.2}$$

where $\Gamma(u) = \mathbb{E}X_0 X_u - \mu(X)^2$ is the autocovariance function. The quantity $\Pi(X)$ is known as the *long-run variance* of the process $X = \{X_t\}_{t\in\mathbb{R}}$. For example, its estimator would naturally enter into the approximate normal confidence intervals for the mean $\mu(X)$. Similarly, define the sample variance

$$\widehat{\sigma}^2_T(X) = \frac{1}{T}\int_0^T (X_s - \overline{X}_T)^2 ds = \frac{1}{T}\int_0^T X_s^2 ds - (\overline{X}_T)^2 = \overline{X^2_T} - (\overline{X}_T)^2 \tag{1.3}$$

and set

$$\begin{aligned}\lim_{T\to\infty} T \operatorname{Var}\left(\begin{pmatrix} \overline{X}_T \\ \overline{X^2_T} \end{pmatrix}\right) &= \begin{pmatrix} \int_{-\infty}^{\infty}\Gamma_X(u)du & \int_{-\infty}^{\infty}\Gamma_{X,X^2}(u)du \\ \int_{-\infty}^{\infty}\Gamma_{X,X^2}(u)du & \int_{-\infty}^{\infty}\Gamma_{X^2}(u)du \end{pmatrix} \\ &=: \begin{pmatrix} \Pi(X) & \Pi(X,X^2) \\ \Pi(X,X^2) & \Pi(X^2) \end{pmatrix} =: \Pi^{(2)}(X,X^2),\end{aligned} \tag{1.4}$$

where $\Gamma_{X,X^2}(u) = \mathbb{E}X(0)X(u)^2 - \mu(X)\mu(X^2)$. One then expects (see Section 3 for details) that

$$\lim_{T\to\infty} T \operatorname{Var}(\widehat{\sigma}^2_T(X)) = \Pi(X^2) - 4\mu(X)\Pi(X,X^2) + 4\mu(X)^2\Pi(X). \tag{1.5}$$

The matrix quantity $\Pi^{(2)}(X,X^2)$ in (1.4) is the *long-run variance* of the vector process $(X,X^2)' = \{(X_t, X_t^2)'\}_{t\in\mathbb{R}}$. Here and throughout, the prime indicates a transpose. The long-run variances are also equal to (possibly up to a multiplicative constant) the spectral densities of the processes at the origin.

Estimation of the long-run variances, such as $\Pi(X)$ and $\Pi^{(2)}(X,X^2)$ above, is a well-studied problem, especially for discrete-time processes (series) X. The often cited seminal reference is Andrews (1991), though the pioneering work in the parallel problem of estimating the spectral density of the process goes back at least to Parzen (1957) and others. We shall refer often to Lu and Park (2014) who focused on continuous-time stochastic processes satisfying a stochastic differential equation and estimation issues from discrete samples.

One of the goals of this work is to revisit estimation of the long-run variances according to the proposed methods, and to assess their performance on the processes associated with oscillatory systems. The case of the inference of the mean is considered in Section 2.1, and that of the variance in Section 3.1. The numerical results are reported in Section 7.1. Aside from some natural adaptations specific to oscillatory systems, this part of the work is rather of the informative nature, especially to those unfamiliar with the relevant literature.

Similarly, we shall also revisit the construction of confidence intervals based on the so-called self-normalization (e.g. Shao (2015)), which does *not* require estimating the long-run variance. Those unfamiliar with the approach might find this surprising but the basic ideas are quite straightforward (see Section 2.2 for the case of the mean, and Section 3.2 for that of the variance). The numerical results for the self-normalization approach and comparison to the approach based on estimating the long-run variances are also reported in Section 7.1. As in other applications of the self-normalization approach, the corresponding confidence intervals tend to be slightly wider on average (compared to those with the estimated long-run variances), but an advantage of the method is its simplicity.

Several other aspects of this work are more novel. Estimation of the long-run variance referenced above (and for that matter the self-normalization approach) assumes implicitly that the long-run variance is non-zero. In Section 4, we also study the case when the long-run variance is zero, which is quite plausible with oscillatory systems and is also the case in the considered data application (Section 7.2). As shown, perhaps somewhat surprisingly, inference about the mean in the case of zero long-run variance is in fact conceptually simpler than that in the case of non-zero long-run variance. In practice, as we argue below, the so-called unit root tests can be used to decide on whether given data are consistent with the hypotheses of non-zero long-run variance.

As another novel aspect of the general problem of constructing confidence intervals, we also study the case when a number of independent records of the same process are available. See Section 5. This is a quite common situation in several applications, for example, related to ship motions, where multiple records would correspond to different experimental runs in an actual model basin, or be obtained by using a computer generation. When considering multiple records, two natural estimators of the long-run variance can be considered: first, the average of the long-run variance estimators of the individual records, and second, the long-run variance estimator deduced from the sample variance of the sample means of individual records. In this regard, we introduce here an estimator that generally outperforms these two natural estimators of the long-run variance, and explain its improved performance. In addition, we propose a way to construct confidence intervals based on self-normalization.

For the sake of clarity, we focus on the case when a sample of the process is available in the continuous time. Some issues behind using discrete samples of the process, which is the more realistic scenario in practice, are discussed in Section 6. Finally, as noted above, Section 7 contains numerical results assessing the performance of the methods described in this work, and Appendix A details a number of processes associated with oscillatory systems that are considered here.

2 Estimation of mean

2.1 Approach involving long-run variance

As in Section 1, consider a continuous-time stationary process $X = \{X_t\}_{t\in\mathbb{R}}$ with mean $\mu(X) = \mathbb{E}X_t$ and finite variance $\sigma^2(X) = \mathrm{Var}(X_t)$. Suppose given a continuous-time sample X_t, $t \in [0, T]$, from which one would like to estimate the mean $\mu(X)$ by using the sample mean $\overline{X}_T$ in (1.1). The approximate normal confidence intervals for $\mu(X)$ would include the long-run variance $\Pi(X)$ in (1.2), which needs to be estimated from data. In defining the estimator of the long-run variance, the so-called kernel function plays a key role.

A kernel K is a function $K : \mathbb{R} \to \mathbb{R}$ such that it is symmetric, $K(0) = 1$ and $\int_{\mathbb{R}} K(x)^2 dx < \infty$.

(In some instances, additional assumptions are made.) Let

$$K_r = \lim_{x\to 0} \frac{1-K(x)}{|x|^r}, \quad r \in \mathbb{Z}^+, \tag{2.1}$$

and $\nu = \max\{r : K_r < \infty\}$. The quantity ν measures smoothness of the kernel $K(x)$ at $x = 0$. Examples of commonly used kernels are Quadratic Spectral (QS): $K(x) = \frac{25}{12\pi^2x^2}\Big(\frac{\sin(6\pi x/5)}{6\pi x/5} - \cos(6\pi x/5)\Big)$, Bartlett: $K(x) = (1-|x|)1_{\{|x|\le 1\}}$, Truncated: $K(x) = 1_{\{|x|\le 1\}}$. (The notation 1_A stands for the indicator function of the set A, that is, $1_A(x) = 1$ if $x \in A$, and $= 0$ otherwise.) For QS kernel, $\nu = 2$ and $K_2 \approx 1.4212$. For Bartlett kernel, $\nu = 1$ and $K_1 = 1$.

Following the main approach found in the literature, the estimator of the long-run variance $\Pi(X)$ in (1.2) is then defined as

$$\widehat{\Pi}_T(X) = \int_{-T}^{T} K\Big(\frac{u}{S_T}\Big)\widehat{\Gamma}_T(u)du, \tag{2.2}$$

where S_T (which is smaller than T) is known as the *bandwidth*, and

$$\widehat{\Gamma}_T(u) = \frac{1}{T}\int_0^{T-u}(X_{s+u} - \overline{X}_T)(X_s - \overline{X}_T)ds, \quad \widehat{\Gamma}_T(-u) = \widehat{\Gamma}_T(u),\ 0 < u < T, \tag{2.3}$$

estimates the autocovariance function of the process X.

The optimal value of the bandwidth, balancing the asymptotic bias and variance of the estimator $\widehat{\Pi}_T(X)$, was derived by Lu and Park (2014), and is given by

$$S_{\mathrm{opt},T} = \Big(\frac{\nu K_\nu^2 C_\nu(X)^2}{\int K(x)^2dx}T\Big)^{1/(2\nu+1)}, \tag{2.4}$$

where K_ν and ν are associated with the kernel K and defined in (2.1) and following (2.1), respectively,

$$C_\nu(X) = \frac{\Lambda_\nu(X)}{\Pi(X)} \tag{2.5}$$

and

$$\Lambda_\nu(X) = \int_{\mathbb{R}} |u|^\nu \Gamma_X(u)du, \quad \nu \in \mathbb{Z}^+. \tag{2.6}$$

Remark 2.1. If S_X is the spectral density of the process X satisfying (by convention)

$$S_X(w) = \int_{-\infty}^{\infty}\Gamma_X(u)\cos(wu)du, \quad \Gamma_X(h) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_X(w)\cos(hw)dw, \tag{2.7}$$

then for even ν,

$$\Lambda_\nu(X) = (-1)^{\nu/2}\frac{d^\nu S_X(w)}{dw^\nu}\Big|_{w=0}.$$

As mentioned above, we also have $\Pi(X) = S_X(0)$.

Note that the constant $C_\nu(X)$ in (2.5) depends on the underlying process X. Several ways of estimating this constant have been suggested. The most popular method seems to be *model-driven*. For this method, a model would be fitted to data and the constant $C_\nu(X)$ would be calculated based on the fitted model. For example, Lu and Park (2014) focus on continuous-time Ornstein-Uhlenbeck (mean reversion) type models. In the context of oscillatory systems, it seems natural to fit to the data a linear oscillator with white noise excitation (see Appendix A). Several ways and references for performing this task from discrete sample are discussed in Section 6.

Another method for estimating $C_\nu(X)$ in (2.5) is *data-driven*. In this method, the idea is to use kernel-based estimators of $\Pi(X)$ (as defined in (2.2) above) and $\Lambda_\nu(X)$ (as in (2.2) but replacing $\widehat{\Gamma}_T(u)$ by $|u|^\nu\widehat{\Gamma}_T(u)$) with some crude choice of the bandwidth S_T. In Section 7 below, we use the bandwidth $S_T = \sqrt{T}$ when estimating $\Pi(X)$ and $\Lambda_\nu(X)$ for the constant $C_\nu(X)$ (but see also Remark 6.2). The model-driven approach is often cited in the literature as being more stable than the data-driven approach, but our simulation study suggests that the data-driven method is often as good, if not superior.

Finally, for the considered estimator $\widehat{\Pi}_T(X)$ of the long-run variance, an approximate normal confidence interval for the mean is defined as

$$\Big(\overline{X}_T - z_\alpha\sqrt{\widehat{\mathrm{Var}}(\overline{X}_T)}, \overline{X}_T + z_\alpha\sqrt{\widehat{\mathrm{Var}}(\overline{X}_T)}\Big), \tag{2.8}$$

where

$$\widehat{\mathrm{Var}}(\overline{X}_T) = \frac{\widehat{\Pi}_T(X)}{T} \tag{2.9}$$

and z_α is the critical value for the standard normal distribution associated with a confidence level α.

2.2 Approach based on self-normalization

The self-normalization approach allows constructing confidence intervals for the mean without estimating the long-run variance. The unfamiliar reader might be surprised that this is possible at all! In fact, there is some "price" to pay: the confidence intervals will *not* be based on a normal distribution and, on average, they tend to be slightly larger. The self-normalization method goes back at least to Lobato (2001) and is summarized nicely in Shao (2015). To explain the approach, we need a slightly stronger result than the central limit theorem, according to which $\sqrt{T}(\overline{X}_T - \mu(X))$ converge to a normal distribution with zero mean and variance $\Pi(X)$ as $T \to \infty$. By the so-called functional central limit theorem, it is expected that, for $z \in [0,1]$,

$$\frac{1}{\sqrt{T}}\int_0^{Tz}(X_s - \mu(X))\,ds \approx \sqrt{\Pi(X)}W(z), \tag{2.10}$$

where $W(z), z \in [0,1]$, is a standard Wiener process (with the term "standard" referring to $\mathbb{E}W(1)^2 = 1$). Note that by setting $z = 1$, (2.10) can also be written as

$$\sqrt{T}(\overline{X}_T - \mu(X)) \approx \sqrt{\Pi(X)}W(1), \tag{2.11}$$

which is just another way to express the central limit theorem.

Now, set

$$N_T^2 = \int_0^1\left(\frac{1}{\sqrt{T}}\int_0^{Tz}\big(X_s - \overline{X}_T\big)\,ds\right)^2 dz \tag{2.12}$$

and note that the quantity inside the square can be expressed as

$$\frac{1}{\sqrt{T}}\int_0^{Tz}(X_s-\mu(X))\,ds - z\sqrt{T}(\overline{X}_T-\mu(X)).$$

The central limit theorems in (2.10) and (2.11) then suggest that

$$N_T^2 \approx \Pi(X)\int_0^1 (W(z)-zW(1))^2\,dz. \tag{2.13}$$

The key observation now is that both (2.11) and (2.13) involve the same long-run variance $\Pi(X)$, which would cancel out by taking the ratio of the two. That is, we expect that

$$\sqrt{T}\frac{\overline{X}_T-\mu(X)}{N_T} \approx \frac{W(1)}{\left(\int_0^1 (W(z)-zW(1))^2\,dz\right)^{1/2}} =: U_0. \tag{2.14}$$

This leads to an approximate confidence interval for the mean given by

$$\left(\overline{X}_T - \frac{u_\alpha N_T}{\sqrt{T}}, \overline{X}_T + \frac{u_\alpha N_T}{\sqrt{T}}\right), \tag{2.15}$$

where u_α is a critical value from the distribution of U_0 for a desired confidence level α. Some critical values for $(U_0)^2$ are tabulated in Lobato (2001). For example, for $\alpha = 90\%$, these yield $u_\alpha = \sqrt{28.31} = 5.32$.

As indicated above, the last point also shows a potential price to pay. The critical values for U_0 are larger than those for the standard normal, potentially leading to larger confidence intervals. This happens only "potentially," since the relationship between $\widehat{\Pi}(X)^{1/2}$ in the normal confidence interval (2.8) and N_T in the confidence interval (2.15) is a priori not clear. In fact, note that the long-run variance $\Pi(X)$ estimated by $\widehat{\Pi}(X)$ is deterministic, whereas N_T is random. A relationship such as $N_T > \widehat{\Pi}(X)^{1/2}$ can only be quantified in probabilistic terms.

3 Estimation of variance

3.1 Approach involving long-run variance

We now turn to inference for the variance $\sigma^2(X)$ of a continuous-time stationary process $X = \{X_t\}_{t\in\mathbb{R}}$, through the sample variance $\widehat{\sigma}_T^2(X)$ defined in (1.3). As indicated around (1.3)–(1.5) in Section 1, the long-run variance $\Pi^{(2)}(X, X^2)$ of the bivariate process $(X, X^2)' = \{(X_t, X_t^2)'\}_{t\in\mathbb{R}}$ in (1.4) now plays a central role. It can be estimated similarly to the long-run variance $\Pi(X)$, as outlined next.

As in Section 2, define the kernel-based estimator of the long-run variance $\Pi^{(2)}(X, X^2)$ as

$$\widehat{\Pi}_T^{(2)}(X, X^2) = \int_{-T}^{T} K\Big(\frac{u}{S_T}\Big)\widehat{\Gamma}_T^{(2)}(u)du, \tag{3.1}$$

where S_T is the *bandwidth*, and

$$\widehat{\Gamma}_T^{(2)}(u) = \frac{1}{T}\int_0^{T-u} \begin{pmatrix} X_{s+u}-\overline{X}_T \\ X_{s+u}^2-\overline{X_T^2}\end{pmatrix}\begin{pmatrix} X_s-\overline{X}_T \\ X_s^2-\overline{X_T^2}\end{pmatrix}' ds, \quad \widehat{\Gamma}_T^{(2)}(-u) = \widehat{\Gamma}_T^{(2)}(u),\ 0<u<T, \tag{3.2}$$

estimates the autocovariance matrix function of the bivariate process $(X, X^2)'$. This also naturally yields the estimators $\widehat{\Pi}_T(X)$, $\widehat{\Pi}_T(X^2)$ and $\widehat{\Pi}_T(X, X^2)$ of the components $\Pi(X)$, $\Pi(X^2)$ and $\Pi(X, X^2)$ of the long-run variance matrix $\Pi^{(2)}(X, X^2)$ in (1.4). The optimal value of the bandwidth, balancing the asymptotic bias and variance of the estimator $\widehat{\Pi}_T^{(2)}(X, X^2)$, was derived by Lu and Park (2014), and is given by

$$S_{\text{opt},T} = \Big(\frac{\nu K_\nu^2 C_\nu(X, X^2)^2}{\int K(x)^2 dx} T\Big)^{1/(2\nu+1)}, \tag{3.3}$$

where K_ν and ν are associated with the kernel K and defined in (2.1) and following (2.1), respectively,

$$C_\nu(X, X^2) = \Big(\frac{\Lambda_\nu^2(X) + \Lambda_\nu^2(X^2)}{\Pi^2(X) + \Pi^2(X^2)}\Big)^{1/2} \tag{3.4}$$

and $\Lambda_\nu(\cdot)$ is defined in (2.6). (In fact, the optimal bandwidth derived by Lu and Park (2014) allows for different weights in balancing the asymptotic bias and variance of the components of the matrix estimator $\widehat{\Pi}_T^{(2)}(X, X^2)$; the bandwidth given above corresponds to equal "diagonal" weights in the weighing scheme.) In practice, the constant $C_\nu(X, X^2)$ is estimated through either the model- or data-driven approaches in the same way as discussed in Section 2 following Remark 2.1.

Finally, with the estimator $\widehat{\Pi}_T^{(2)}(X, X^2) =: (\widehat{\Pi}_T(X)\ \widehat{\Pi}_T(X, X^2); \widehat{\Pi}_T(X, X^2)\ \widehat{\Pi}_T(X^2))$ of the long-run variance matrix given by (3.1)–(3.2), an approximate normal confidence interval for the variance is defined as

$$\Big(\widehat{\sigma}_T^2(X) - z_\alpha \sqrt{\widehat{\mathrm{Var}}(\widehat{\sigma}_T^2(X))}, \widehat{\sigma}_T^2(X) + z_\alpha \sqrt{\widehat{\mathrm{Var}}(\widehat{\sigma}_T^2(X))}\Big), \tag{3.5}$$

where

$$\widehat{\mathrm{Var}}(\widehat{\sigma}_T^2(X)) = \frac{\widehat{\Pi}_T(X^2) - 4\overline{X}_T \widehat{\Pi}_T(X, X^2) + 4(\overline{X}_T)^2 \widehat{\Pi}_T(X)}{T} \tag{3.6}$$

and z_α is the critical value for the standard normal distribution associated with a confidence level α. The numerator in (3.6) is motivated by the form of the asymptotic variance of $\widehat{\sigma}_T^2(X)$ in (1.5). The latter form itself is a consequence of the delta method applied to the right-hand side of (1.3) and using (1.4).

Remark 3.1. A similar application of the delta method to the sample standard deviation $\widehat{\sigma}_T(X) = (\widehat{\sigma}_T^2(X))^{1/2}$ yields an approximate normal confidence interval for the standard deviation $\sigma(X)$ of the process, defined as

$$\Big(\widehat{\sigma}_T(X) - z_\alpha \sqrt{\widehat{\mathrm{Var}}(\widehat{\sigma}_T(X))}, \widehat{\sigma}_T(X) + z_\alpha \sqrt{\widehat{\mathrm{Var}}(\widehat{\sigma}_T(X))}\Big), \tag{3.7}$$

where z_α is as in (3.5) and

$$\widehat{\mathrm{Var}}(\widehat{\sigma}_T(X)) = \frac{\widehat{\Pi}_T(X^2) - 2(\overline{X}_T)^2 \widehat{\Pi}_T(X) - 4(\overline{X}_T)^2 \widehat{\Pi}_T(X, X^2)}{T(4(\overline{X^2}_T - (\overline{X}_T)^2))}. \tag{3.8}$$

Remark 3.2. The inference approach outlined above can also be applied to estimating quantities of interest other than the mean, the variance or the standard deviation. For example, the autocovariance of a process X at lag $h > 0$, defined as

$$\gamma_X(h) = \mathbb{E}X_0 X_h - \mu(X)^2 = \mathbb{E}(X_0 - \mu(X))(X_h - \mu(X)),$$

is naturally estimated through

$$\widehat{\gamma}_T(h) = (\overline{XX}_{\cdot+h})_T - (\overline{X}_T)^2,$$

where

$$(\overline{XX}_{\cdot+h})_T = \frac{1}{T}\int_0^{T-h}(X_{s+h} - \overline{X}_T)(X_s - \overline{X}_T)^2 ds.$$

In this case, the long-run variance matrix of the bivariate process $(X, \overline{XX}_{\cdot+h})' = \{(X_t, X_t X_{t+h})',\ t \in \mathbb{R}\}$ is relevant and could be estimated similarly to (3.1).

3.2 Approach based on self-normalization

A simple-minded approach would be to treat the sample variance in (1.3) just as the sample mean of $(X_s - \overline{X}_T)^2$ or, after an approximation, of $(X_s - \mathbb{E}X_s)^2$, and then just apply the approaches outlined in Section 2.2 with X_s replaced by $(X_s - \overline{X}_T)^2$. This effectively assumes that in the self-normalization approach, working with $(X_s - \overline{X}_T)^2$ is the same, in the asymptotic sense for large T, as working with $(X_s - \mathbb{E}X_s)^2$. Some justification for this in a similar setting can be found in Lobato (2001), Sec. 3.

4 Degenerate case

The discussion in Sections 2 and 3 assumes implicitly that the long-run variance $\Pi(X) \neq 0$. What happens in the case

$$\Pi(X) = 0 \quad (\text{equivalently}, S_X(0) = 0), \tag{4.1}$$

where S_X is the spectral density of the process X defined in Remark 2.1? This case is known as *degenerate* (e.g. Lee (2010)). Having zero long-run variance is quite plausible for processes associated with oscillatory systems, since their autocovariance function exhibits naturally an oscillating pattern and thus can integrate to 0.

In the degenerate case, since $S_X(0) = 0$ and $S_X(w)$ is even and often smooth at $w = 0$, it is expected that

$$\int_{-\infty}^{\infty} \frac{S_X(w)}{|w|^2} dw < \infty. \tag{4.2}$$

Under this assumption, one can write

$$\int_0^t (X_s - \mu(X))ds = V_t - V_0, \tag{4.3}$$

where $\{V_t\}_{t\in\mathbb{R}}$ is a *stationary* process with zero mean. Indeed, by writing X_s in the spectral domain as $X_s - \mu(X) = \int_{-\infty}^{\infty} e^{isw}\sqrt{S_X(w)}Z(dw)$ with a complex-valued random measure $Z(dw)$ having orthogonal increments and $\mathbb{E}|Z(dw)|^2 = dw/(2\pi)$, note that

$$\int_0^t (X_s - \mu(X))ds = \int_{-\infty}^{\infty} \Big(\int_0^t e^{isw} ds\Big)\sqrt{S_X(w)}Z(dw)$$

$$= \int_{-\infty}^{\infty} e^{itw}\frac{\sqrt{S_X(w)}}{iw}Z(dw) - \int_{-\infty}^{\infty}\frac{\sqrt{S_X(w)}}{iw}Z(dw),$$

that is, the stationary process V_t can be taken as $\int_{-\infty}^{\infty} e^{itw}(\sqrt{S_X(w)}/iw)Z(dw)$ and has the spectral density $S_X(w)/|w|^2$. It follows from (4.3) that

$$T(\overline{X}_T - \mu(X)) = \int_0^T (X_s - \mu(X))ds \stackrel{d}{=} V_T - V_0. \tag{4.4}$$

Note that, compared to (1.2), the convergence rate of $\mathrm{Var}(\overline{X}_T)$ is then T^2, and not T. Note also that $V_T - V_0$ does not need to be Gaussian.

Thus, in view of (4.3), the integrated process $\int_0^t (X_s - \mu(X))ds$ is nearly stationary itself, especially as t increases (since dependence between V_t and V_0 is then expected negligible). In particular, we expect that, for large T,

$$T(\overline{X}_T - \mu(X)) \approx V_0' - V_0, \tag{4.5}$$

where V_0' is an independent copy of V_0, and an approximate (not necessarily normal) confidence interval can be constructed as

$$\Big(\overline{X}_T - \frac{t_\alpha(\widehat{V_0' - V_0})}{T}, \overline{X}_T + \frac{t_\alpha(\widehat{V_0' - V_0})}{T}\Big), \tag{4.6}$$

where $t_\alpha(V_0' - V_0)$ denotes the $(1+\alpha)/2$ quantile of $V_0' - V_0$ and the hat its estimator. Note that, in practice, the value $(-V_0)$ can be estimated as the sample mean of $\int_0^t (X_s - \overline{X}_T)ds$, $0 < t < T$, and then removed from $V_t - V_0$ to obtain a realization of V_t. The latter realization can then be used to estimate the corresponding quantile of $V_0' - V_0$: in practice, we use resampling from a realization of V_t to obtain a sample of the values of $V_0' - V_0$, from which we then select a desired quantile.

We also note that from a practical perspective, a decision needs to be made on whether the underlying process is in the degenerate case or, equivalently, whether the integrated process $\int_0^t (X_s - \mu(X))ds$ can be viewed stationary. As discussed in Section 7 below, this can be achieved by using one of the available so-called unit root tests.

Finally, the discussion above concerns the case of estimating the mean. Turning to estimation of the variance, recall that it involves estimation of the mean of the square process X^2. If X can be expected to fall in the degenerate case, we do not expect this to be the case with the square process X^2. For example, if X is Gaussian, then $\Gamma_{X^2}(h) = 2(\Gamma_X(h))^2$ and $\Pi(X^2) = \int_{-\infty}^{\infty} \Gamma_{X^2}(h) \neq 0$. If X falls in the degenerate case but X^2 does not, then

$$\lim_{T\to\infty} T\,\mathrm{Var}(\widehat{\sigma}_T^2(X)) = \Pi(X^2),$$

since the rate of convergence of $(\overline{X}_T)^2$ to $\mu(X)^2$ is faster than $\sqrt{T}$. Hence, only the long-run variance of the process X^2 needs to be estimated.

5 The case of multiple records

5.1 Approach involving long-run variance

Consider a number of records $X_t^{(r)}$, $t \in [0, T_r]$, $r = 1, \ldots, R$, that are independent across r. Suppose that

$$T_r = C_r T \quad \text{with} \quad C_r \in (0,1),\ \sum_r C_r = 1. \tag{5.1}$$

Thus, $\sum_r T_r = T$. For example, the case of the records of equal length corresponds to

$$C_r = 1/R \quad \text{and} \quad T_r = T/R.$$

Considering inference for the mean first, it can be estimated through

$$\overline{X}_T = \sum_r C_r \overline{X}^{(r)}_{T_r}, \tag{5.2}$$

that is, as the appropriately weighted average of the sample means of the R records. Note that

$$T \operatorname{Var}(\overline{X}_T) = T \sum_r C_r^2 \operatorname{Var}(\overline{X}^{(r)}_{T_r}) = \sum_r C_r T_r \operatorname{Var}(\overline{X}^{(r)}_{T_r}) \to \sum_r C_r \Pi(X) = \Pi(X). \tag{5.3}$$

For confidence intervals, the long-run variance $\Pi(X)$ thus again needs to be estimated, but this time from R records. We focus in this section on the non-degenerate case when the long-run variance is non-zero.

We note first that several natural estimators of the long-run variance can be introduced in the case of multiple records. First, there is the *weighted average* of the estimators of the long-run variance across the records, defined as

$$\widehat{\Pi}_{\text{ave},T}(X) = \sum_r C_r \widehat{\Pi}^{(r)}_{T_r}(X) = \sum_r C_r \int_{-T_r}^{T_r} K\Big(\frac{u}{S_{T_r}}\Big) \widehat{\Gamma}^{(r)}_{T_r}(u) du, \tag{5.4}$$

where $\widehat{\Pi}^{(r)}_{T_r}(X)$ is the estimator of the long-run variance for the rth record. Second, a *direct estimator* of the long-run variance can be defined as a properly normalized sample variance of the sample means across the records. Indeed, consider the case of equal length records and recall that $\Pi(X)$ approximately equals $T_r \operatorname{Var}(\overline{X}^{(r)}_{T_r})$. But since one now has R sample means $\overline{X}^{(r)}_{T_r}$, $r = 1, \ldots, R$, one can naturally estimate $\operatorname{Var}(\overline{X}^{(r)}_{T_r})$ through the sample variance of $\overline{X}^{(r)}_{T_r}$, $r = 1, \ldots, R$, and then normalize it by T_r to get the estimator of the long-run variance. In the case of equal length records, this direct estimator is

$$\widehat{\Pi}_{\text{dir},T}(X) = \frac{T_r}{R} \sum_r \Big(\overline{X}^{(r)}_{T_r} - \overline{X}_T\Big)^2 = T \sum_r \frac{1}{R^2} \Big(\overline{X}^{(r)}_{T_r} - \overline{X}_T\Big)^2. \tag{5.5}$$

This motivates the following definition in the general case,

$$\widehat{\Pi}_{\text{dir},T}(X) = T \sum_r C_r^2 \Big(\overline{X}^{(r)}_{T_r} - \overline{X}_T\Big)^2. \tag{5.6}$$

Note that this definition is also consistent with the expressions in (5.3) where $\operatorname{Var}(\overline{X}^{(r)}_{T_r})$ is replaced by a "naive" estimator $(\overline{X}^{(r)}_{T_r} - \overline{X}_T)^2$. Note also that the estimator $\widehat{\Pi}_{\text{dir},T}(X)$ is unique to the case of multiple records.

Which of the estimators, $\widehat{\Pi}_{\text{ave},T}(X)$ or $\widehat{\Pi}_{\text{dir},T}(X)$, should be preferred? The simulation results in Section 7 suggest that $\widehat{\Pi}_{\text{ave},T}(X)$ is superior to $\widehat{\Pi}_{\text{dir},T}(X)$ in terms of the variance, but that it can also be inferior in terms of the bias. In fact, another estimator can be introduced which enjoys both advantages of the two natural estimators. To motivate the definition of the new estimator, we shall rewrite the estimator $\widehat{\Pi}_{\text{dir},T}(X)$ in a different form as follows.

Observe that

$$\widehat{\Pi}_{\mathrm{dir},T}(X) = T\sum_r C_r^2\Big(\frac{1}{T_r}\int_0^{T_r} X_s^{(r)}ds - \overline{X}_T\Big)^2 = T\sum_r \frac{C_r^2}{T_r^2}\Big(\int_0^{T_r}(X_s^{(r)} - \overline{X}_T)ds\Big)^2$$

$$= \sum_r \frac{C_r}{T_r}\int_0^{T_r}\int_0^{T_r}(X_s^{(r)} - \overline{X}_T)(X_t^{(r)} - \overline{X}_T)dsdt = \sum_r C_r\int_{-T_r}^{T_r}\widehat{\Gamma}_{0,T_r}^{(r)}(u)du, \tag{5.7}$$

where

$$\widehat{\Gamma}_{0,T_r}^{(r)}(u) = \frac{1}{T_r}\int_0^{T_r-u}(X_{s+u}^{(r)} - \overline{X}_T)(X_s^{(r)} - \overline{X}_T)ds, \quad \widehat{\Gamma}_{0,T_r}^{(r)}(-u) = \widehat{\Gamma}_{0,T_r}^{(r)}(u),\ 0 < u < T_r. \tag{5.8}$$

There are two key differences between the average estimator (5.4) and the direct estimator (5.7): first, the average estimator uses a kernel function to down-weigh the effects of the estimator of the autocovariance function at large lags (and thus to reduce the variance), and second, the direct estimator employs the mean across the records in estimating the autocovariance function (and thus reducing the bias). In view of these differences, it natural to introduce the following estimator of the long-run variance,

$$\widehat{\Pi}_T(X) = \sum_r C_r\int_{-T_r}^{T_r} K\Big(\frac{u}{S_{T_r}}\Big)\widehat{\Gamma}_{0,T_r}^{(r)}(u)du =: \sum_r C_r\widehat{\Pi}_{0,T_r}^{(r)}(X), \tag{5.9}$$

that is, defined as the average of the estimators of the long-run variance for the R records which use the mean across all records in the estimation of the autocovariance function. As shown in Section 7, this estimator generally outperforms the average and direct estimators of the long-run variance.

Finally, we note that the superiority of the estimator in (5.9) is more apparent when X is replaced by the square process X^2, which is relevant when estimating the variance $\sigma^2(X)$ of the process (Section 3 above).

5.2 Approach based on self-normalization

This section describes one possible and natural way to adapt the self-normalization approach to the case of multiple records. Note that the sample mean $\overline{X}_T$ across the R records in (5.2) can also be expressed as

$$\overline{X}_T = \frac{1}{T}\int_0^T X_s ds, \tag{5.10}$$

where $T = \sum_r T_r$ and

$$X_s = \sum_r X^{(r)}_{s-\sum_{q=1}^{r-1}T_q}1_{[\sum_{q=1}^{r-1}T_q, \sum_{q=1}^{r}T_q)}(s) \tag{5.11}$$

is the process obtained by "concatenating" the R records together. As in (2.12), consider the normalization quantity for the process (5.10)–(5.11) defined as

$$N_T^2 = \int_0^1\left(\frac{1}{\sqrt{T}}\int_0^{Tz}\left(X_s - \overline{X}_T\right)ds\right)^2 dz. \tag{5.12}$$

$R \setminus p$	0.95	0.97	0.99	0.995
1	5.2836	6.7710	8.6115	9.9781
2	5.3116	6.7432	8.6533	10.038
3	5.3626	6.7414	8.6497	10.120
4	5.3224	6.7806	8.5813	10.031
5	5.2818	6.6775	8.6066	10.061
10	5.3057	6.7258	8.5940	10.147

Table 1: Quantiles of $U_{0,R}$.

By using the relations analogous to (2.10) but for the R records, one can show that

$$N_T^2 \approx \sum_{r=1}^{R} \int_0^1 \left(\sum_{q=1}^{r-1} \sqrt{C_q} W^{(q)}(1) + \sqrt{C_r} W^{(r)}(w) - (\overline{C}_{r-1} + C_r w) \sum_{q=1}^{R} \sqrt{C_q} W^{(q)}(1) \right)^2 dw, \quad (5.13)$$

where $\overline{C}_q = \sum_{i=1}^{q} C_i$ and $\{W^{(1)}(w)\}_{w\in[0,1]}, ..., \{W^{(R)}(w)\}_{w\in[0,1]}$ are R independent standard Brownian motions.

For example, in the case of R records of equal length with $C_r = 1/R, r = 1, ..., R$ the above discussion suggests that

$$\sqrt{T}\frac{\overline{X}_T - \mu(X)}{N_T} \approx \frac{\sum_{r=1}^{R} W^{(r)}(1)}{\left(\sum_{r=1}^{R} \int_0^1 \left(\sum_{q=1}^{r-1} W^{(q)}(1) + W^{(r)}(w) - \frac{r-1+w}{R} \sum_{q=1}^{R} W^{(q)}(1) \right)^2 \frac{1}{R} dw \right)^{1/2}} =: U_{0,R}. \quad (5.14)$$

Several quantiles of the distribution of $U_{0,R}$ for several values of R are tabulated in Table 1. They were computed as the suitable quantiles obtained from multiple generated values of $U_{0,R}$, calculated after discretizing the intergal in the last denominator in (5.14). More specifically, a random walk $W(w)$ was approximated by $N^{-1/2} \sum_{n=1}^{[Nw]} Z_n$ with i.i.d. standard normal Z_n and $N = 10,000$, and $100,000$ independent replications of $U_{0,R}$ were used.

With the quantiles available from Table 1, a confidence interval for the mean based on multiple records can be constructed as in (2.15).

6 Discretization and other issues

In Sections 2–5, we assumed that a continuous-time sample (or multiple samples) of the analyzed process is available. In practice, however, a discrete(-time) sample of the process is given, namely, $X_{k\Delta}$, $k = 1, \ldots, T/\Delta$, where $\Delta > 0$ is the discretization step and we assume for simplicity that

$$\frac{T}{\Delta} = n \quad (6.1)$$

is an integer. We examine here a number of issues behind using a discrete sample rather than a continuous-time one, when estimating the process mean and the process variance. Our goal is not to provide any formal proofs (as e.g. in Lu and Park (2014) concerning estimation of the long-run variance through discrete sample) but rather to guide a practitioner through a number of issues that arise from the practical perspective. This section concerns and focuses on the approach based on the long-run variance, which is the more sophisticated approach considered in this paper.

6.1 Discretizing proposed estimators

The various estimators of long-run variances introduced above (see (2.2), (3.1), (5.4), (5.6) and (5.9)) involve integrals in continuous time which can be discretized when working with discrete samples. For example, the discrete version of the estimator (2.2) of the long-run variance is defined as

$$\begin{aligned}\widehat{\Pi}_{T,\Delta}(X) &= \sum_{j=-(T-\Delta)/\Delta}^{(T-\Delta)/\Delta} K\Big(\frac{j\Delta}{S_{T,\Delta}}\Big)\widehat{\Gamma}_{T,\Delta}(j\Delta)\Delta \\ &= \Delta \sum_{j=-(n-1)}^{n-1} K\Big(\frac{j}{m}\Big)\widehat{\Gamma}_n(j) =: \Delta\cdot\Omega_n, \end{aligned} \tag{6.2}$$

where

$$m = \frac{S_{T,\Delta}}{\Delta} \tag{6.3}$$

is referred to as the bandwidth for the discrete sample with some discrete version $S_{T,\Delta}$ of the continuous-time bandwidth S_T,

$$\begin{aligned}\widehat{\Gamma}_{T,\Delta}(j\Delta) &= \frac{1}{T}\sum_{k=1}^{(T-j\Delta)/\Delta} (X_{(k+j)\Delta} - \overline{X}_{T,\Delta})(X_{k\Delta} - \overline{X}_{T,\Delta})\Delta \\ &= \frac{1}{n}\sum_{k=1}^{n-j}(X_{(k+j)\Delta} - \overline{X}_{T,\Delta})(X_{k\Delta} - \overline{X}_{T,\Delta}) \\ &= :\widehat{\Gamma}_n(j), \quad j = 0,1,\ldots,n-1, \end{aligned} \tag{6.4}$$

and $\widehat{\Gamma}_{T,\Delta}(-j\Delta) = \widehat{\Gamma}_{T,\Delta}(j\Delta)$, $\widehat{\Gamma}_n(-j) = \widehat{\Gamma}_n(j)$, and

$$\overline{X}_{T,\Delta} = \frac{1}{T}\sum_{k=1}^{T/\Delta} X_{k\Delta}\Delta = \frac{1}{n}\sum_{k=1}^{n} X_{k\Delta}. \tag{6.5}$$

The expressions following the first equality signs in (6.2), (6.4) and (6.5) are written as to emphasize that the integrals in the corresponding continuous-time estimators are discretized. The quantity Ω_n in (6.2) is the usual estimator of the long-run variance

$$\Pi(X_\Delta) = \sum_{k=-\infty}^{\infty} \Gamma_{X_\Delta}(k) \quad \text{with } \Gamma_{X_\Delta}(k) = \mathbb{E}X_0X_{k\Delta} - \mu(X)^2 = \Gamma_X(k\Delta),$$

of the discrete time series $X_\Delta = \{X_{k\Delta}\}_{k\in\mathbb{Z}}$. Note that it is meaningful to multiply Ω_n by Δ in (6.2) when estimating the long-run variance $\Pi(X)$ of the continuous time process $X = \{X_T\}_{t\in\mathbb{R}}$ since

$$\Delta\cdot\Pi(X_\Delta) = \sum_{k=-\infty}^{\infty}\Gamma_X(k\Delta)\Delta \approx \int_{\mathbb{R}}\Gamma_X(u)du = \Pi(X). \tag{6.6}$$

The discrete versions of other introduced estimators, namely, (3.1), (5.4), (5.6) and (5.9), are defined in an analogous fashion by discretizing all the integrals involved.

6.2 Model-driven bandwidth selection

The discrete version $\widehat{\Pi}_{T,\Delta}$ of the estimator of the long-run variance defined in (6.2) uses the bandwidth m in (6.3), which requires a discrete version of $S_{T,\Delta}$ of the continuous-time bandwidth S_T defined in (2.4). We focus here on the optimal continuous-time bandwidth $S_{\text{opt},T}$ in (2.4). In Section 2, we discussed two ways of computing $S_{\text{opt},T}$ and, more specifically, the constant $C_\nu(X)$ entering $S_{\text{opt},T}$: the data-driven and the model-driven approaches. The data-driven approach for discrete sample is discussed in Section 6.3 below. For the model-driven approach, it was suggested to use a linear oscillator with white noise excitation (see Appendix A) as an underlying model. The question then is how to fit such continuous-time model given a discrete sample.

The problem of fitting a linear oscillator with white noise excitation from discrete sample has been studied quite extensively in the literature (in fact, not just for a linear oscillator but for the more general class of the so-called continuous AR models). See e.g. Soderstrom et al. (1997), Fan et al. (1999), Kirshner et al. (2014), Lin and Lototsky (2011; 2014), Pham (2000). As discussed in these references, there are delicate issues in how a discrete version of a linear oscillator is formulated and fitted to the data (assuming the linear oscillator is indeed the underlying model). For example, a scheme that works is to use the following discretization of the derivatives at $t = k\Delta$,

$$\begin{aligned} \ddot{X}_t &: \frac{X_{t+3\Delta} - 2X_{t+2\Delta} + X_{t+\Delta}}{\Delta^2} =: (\Delta^2 X)_t, \\ \dot{X}_t &: \frac{X_{t+\Delta} - X_t}{\Delta} =: (\Delta X)_t, \end{aligned} \tag{6.7}$$

and to estimate the coefficients of the model by regressing $(\Delta^2 X)_t$ on $-(\Delta X)_t$ and $-X_t$. See Soderstrom et al. (1997), Example 3.3 on page 662. An estimator of the coefficient σ controlling the strength of the white noise excitation (see Appendix A) can also be given. Other "valid" discretization schemes are available as well, but in our simulations, we use the discretization scheme (6.7) only.

6.3 Connections to discrete time analysis

We noted following the relation (6.5) that the quantity Ω_n in (6.2) is a commonly used estimator of the long-run variance $\Pi(X_\Delta)$ of the discrete-time series $X_\Delta = \{X_{k\Delta}\}_{k\in\mathbb{Z}}$. It is naturally multiplied by Δ in (6.2) in view of (6.6). From the perspective of confidence intervals (2.8)–(2.9), note that the standard error of the sample mean used in the intervals is

$$\sqrt{\frac{\widehat{\Pi}_{T,\Delta}(X)}{T}} = \sqrt{\frac{\Delta \cdot \Omega_n}{T}} = \sqrt{\frac{\Omega_n}{n}}, \tag{6.8}$$

which is exactly the same if the whole analysis is carried out for the discrete sample itself, without any reference to the continuous-time process.

Despite the latter conclusion, however, a more delicate issue concerns the choice of the bandwidth m when used with the discrete sample $X_{k\Delta}$, $k = 1, \ldots, n$. The bandwidth m was defined in (6.3) by relating it to a discrete version $S_{T,\Delta}$ of the continuous-time bandwidth S_T, but it can also be defined with the reference to the underlying discrete-time series $X_\Delta = \{X_{k\Delta}\}_{k\in\mathbb{Z}}$ alone. For a discrete-time series X_Δ, the optimal bandwidth is defined as (Lu and Park (2014))

$$m_{\text{opt}} = \Big(\frac{\nu K_\nu^2 C_\nu(X_\Delta)^2}{\int K(x)^2 dx} n\Big)^{1/(2\nu+1)}, \tag{6.9}$$

where ν and K_ν are associated with a kernel function K as in Section 2,

$$C_\nu(X_\Delta) = \frac{\Lambda_\nu(X_\Delta)}{\Pi(X_\Delta)} \tag{6.10}$$

and

$$\Lambda_\nu(X_\Delta) = \sum_{k=-\infty}^{\infty} |k|^\nu \Gamma_{X_\Delta}(k), \quad \Pi(X_\Delta) = \sum_{k=-\infty}^{\infty} \Gamma_{X_\Delta}(k) \tag{6.11}$$

with $\Gamma_{X_\Delta}(h) = \mathbb{E}X_0 X_{h\Delta} - \mu(X)^2$. It is not immediate to see but the two optimal bandwidths (6.9) and (2.4) are, in fact, connected in a natural way, as we explain next.

6.3.1 The case of the data-driven approach

To understand the relationship between the optimal bandwidths (6.9) and (2.4), consider first the case of the data-driven approach. In this approach and with discrete sample, the constant $C_\nu(X_\Delta)$ in (6.10) is estimated as

$$\widehat{C}_\nu(X_\Delta) = \frac{\widehat{\Lambda}_\nu(X_\Delta)}{\widehat{\Pi}(X_\Delta)}, \tag{6.12}$$

where

$$\widehat{\Pi}(X_\Delta) = \sum_{j=-(n-1)}^{n-1} K\Big(\frac{j}{m_0}\Big)\Gamma_n(j), \quad \widehat{\Lambda}_\nu(X_\Delta) = \sum_{j=-(n-1)}^{n-1} K\Big(\frac{j}{m_0}\Big)|j|^\nu \Gamma_n(j) \tag{6.13}$$

and m_0 is some preliminary crude estimate of the bandwidth. But note that

$$\Delta \cdot \widehat{\Pi}(X_\Delta) = \sum_{j=-(T-\Delta)/\Delta}^{(T-\Delta)/\Delta} K\Big(\frac{j\Delta}{m_0\Delta}\Big)\Gamma_T(j\Delta)\Delta =: \widehat{\Pi}_{T,\Delta}(X), \tag{6.14}$$

$$\Delta^{\nu+1} \cdot \widehat{\Lambda}_\nu(X_\Delta) = \sum_{j=-(T-\Delta)/\Delta}^{(T-\Delta)/\Delta} K\Big(\frac{j\Delta}{m_0\Delta}\Big)|j\Delta|^\nu \Gamma_T(j\Delta)\Delta =: \widehat{\Lambda}_{\nu,T,\Delta}(X). \tag{6.15}$$

The quantities $\widehat{\Pi}_{T,\Delta}(X)$ and $\widehat{\Lambda}_{T,\Delta}(X)$ are the discrete-time estimators of the constants $\Pi(X)$ and $\Lambda_\nu(X)$ when using the bandwidth $S_{0,T} = m_0\Delta$, and would similarly appear in the data-driven discrete-time estimator of the constant $C_\nu(X)$ in (2.5), that is,

$$\widehat{C}_{\nu,\Delta}(X) = \frac{\widehat{\Lambda}_{\nu,T,\Delta}(X)}{\widehat{\Pi}_{T,\Delta}(X)}. \tag{6.16}$$

The relations (6.14) and (6.15) show that

$$\widehat{C}_{\nu,\Delta}(X) = \Delta^{\nu+\frac{1}{2}}\widehat{C}_\nu(X_\Delta) \tag{6.17}$$

as long us the underlying bandwidths satisfy

$$S_{0,T} = m_0\Delta. \tag{6.18}$$

The relation (6.17) then implies that (the discrete estimator of) the optimal bandwidth $S_{\mathrm{opt},T}$ in (2.4) and (the estimator of) the optimal bandwidth m_{opt} in (6.9) are related as

$$S_{\mathrm{opt},T} = m_{\mathrm{opt}}\Delta. \tag{6.19}$$

Thus, in view of (6.8) and in the case of the data-driven approach, when the optimal bandwidths are used with the underlying bandwidths satisfying (6.18), inference about the sample mean is exactly the same when applying the discrete time series analysis to the series X_Δ and when the analysis is based on the assumption of the underlying continuous-time process X.

6.3.2 The case of the model-driven approach

Turning to the relationship between the optimal bandwidths (6.9) and (2.4) in the model-driven approach, the situation is more delicate and we shall make just a few points through a concrete example.

For discrete-time series, commonly chosen underlying models for the bandwidth calculation are the AR series (Andrews (1991), Lu and Park (2014)). For example, for the AR(2) series $Y = \{Y_n\}_{n\in\mathbb{Z}}$ satisfying $Y_n - \phi_1 Y_{n-1} - \phi_2 Y_{n-2} = \sigma_Z Z_n$ with white noise series $\{Z_n\}$ and coefficients ϕ_1, ϕ_2, σ_Z, and when $\nu = 2$, one has

$$\Pi(Y) = \frac{\sigma_Z^2}{2\pi}\frac{1}{(1-\phi_1-\phi_2)^2}, \quad \Lambda_2(Y) = \frac{\sigma_Z^2}{2\pi}\frac{8\phi_2 + 2\phi_1 - 2\phi_1\phi_2}{(1-\phi_1-\phi_2)^4}$$

(e.g. Andrews (1991)). Then,

$$C_2(Y) = \frac{8\phi_2 + 2\phi_1 - 2\phi_1\phi_2}{(1-\phi_1-\phi_2)^2} \tag{6.20}$$

and in the model-driven approach, this constant would be estimated as

$$\widehat{C}_2(Y) = \frac{8\widehat{\phi}_2 + 2\widehat{\phi}_1 - 2\widehat{\phi}_1\widehat{\phi}_2}{(1-\widehat{\phi}_1-\widehat{\phi}_2)^2}, \tag{6.21}$$

where $\widehat{\phi}_1, \widehat{\phi}_2$ are the AR(2) parameters estimated from a discrete sample. Since the AR(2) series can exhibit oscillatory behavior, this series also seems natural to consider for discrete-time series $Y = X_\Delta$.

On the other hand, there is an analogous constant $C_2(X)$ for the model-driven approach in the continuous time supposing a linear oscillator with white noise excitation, defined by (A.1) and (A.2) in Appendix A and characterized by the parameters δ, w_0^2 and σ. For this model (see Appendix A),

$$\Pi(X) = \frac{\sigma^2}{w_0^4}, \quad \Lambda_2(X) = \frac{4(2\delta^2 - w_0^2)\sigma^2}{w_0^8},$$

so that

$$C_2(X) = \frac{4(2\delta^2 - w_0^2)}{w_0^4}. \tag{6.22}$$

What is the relationship between the constant $\widehat{C}_2(Y)$ in (6.21) when $Y = X_\Delta$ and the constant $C_2(X)$, assuming that a linear oscillator with the white noise excitation is the underlying continuous-time model?

To answer this question, let $\widehat{\phi}_1, \widehat{\phi}_2$ be the AR(2) coefficients estimated from the sample $X_{k\Delta}$, $k = 1, \ldots, n$. These are the regression coefficients when regressing $X_{(k+2)\Delta}$ on $X_{(k+1)\Delta}$ and $X_{k\Delta}$, respectively. On the other hand, let $-2\widehat{\delta}_1$ and $-\widehat{w}_{1,0}^2$ be the regression coefficients when regressing $(X_{(k+2)\Delta} - 2X_{(k+1)\Delta} + X_{k\Delta})/\Delta^2$ (which can be thought as a discrete version of $\ddot{X}_{k\Delta}$) on $(X_{(k+1)\Delta} - X_{k\Delta})/\Delta$ (which can be thought as a discrete version of $\dot{X}_{k\Delta}$) and $X_{k\Delta}$. By comparing the two regressions, note that

$$\widehat{\phi}_1 = 2 - 2\widehat{\delta}_1\Delta, \quad \widehat{\phi}_2 = -1 + 2\widehat{\delta}_1\Delta - \widehat{w}_{1,0}^2\Delta^2. \tag{6.23}$$

Then, substituting these expressions into (6.21), we obtain that

$$\widehat{C}_2(X_\Delta) = \frac{4(2\widehat{\delta}_1^2 - \widehat{w}_{1,0}^2 - \widehat{\delta}_1\widehat{w}_{1,0}^2\Delta)}{\widehat{w}_{1,0}^4} \tag{6.24}$$

It is known (see Soderstrom et al. (1997), Example 3.3 on page 662) that

$$\widehat{\delta}_1 \xrightarrow{p} \frac{2}{3}\delta, \quad \widehat{w}_{1,0}^2 \xrightarrow{p} w_0^2, \tag{6.25}$$

as $\Delta \to 0$, where $\xrightarrow{p}$ denotes the convergence in probability. Thus,

$$\widehat{C}_2(X_\Delta) \xrightarrow{p} \frac{4(2(2/3)^2\delta^2 - w_0^2)}{w_0^4} \neq \frac{4(2\delta^2 - w_0^2)}{w_0^4} = C_2(X). \tag{6.26}$$

That is, the two constants in the limit of $\Delta \to 0$ are slightly different, and the two analyses, the discrete-time series analysis for the series X_Δ and the continuous-time process analysis for X, would give slightly different results in the model-driven approach.

Remark 6.1. In the regression above, we used the discretization scheme

$$\begin{aligned} \ddot{X}_t &: \frac{X_{t+2\Delta} - 2X_{t+\Delta} + X_t}{\Delta^2}, \\ \dot{X}_t &: \frac{X_{t+\Delta} - X_t}{\Delta}, \end{aligned} \tag{6.27}$$

for the underlying oscillator but it did not lead to the consistent estimators as noted in (6.25). A discretization scheme leading to consistent estimators was given in (6.7). Since this scheme involves $X_{t+3\Delta}$, $X_{t+2\Delta}$, $X_{t+\Delta}$ and X_t, it may appear to correspond to fitting an AR(3) series to the discrete sample $X_{k\Delta}$, $k = 1, \ldots, n$. But note that the regression of $(\Delta^2 X)_t$ on $(-\Delta X)_t$ and $(-X_t)$ for $t = k\Delta$ has the coefficient equal exactly to 2 at $X_{t+2\Delta}$ since $X_{t+2\Delta}$ appears with a factor of (-2) in $(\Delta^2 X)_t$ and not in $(-\Delta X)_t$, nor in $(-X_t)$. Thus, using the consistent discretization scheme would not be equivalent to fitting the AR(3) series to the discrete sample $X_{k\Delta}$, $k = 1, \ldots, n$.

6.4 Range of discretization step Δ

As the discretization step Δ approaches 0, the discrete version $\widehat{\Pi}_{T,\Delta}$ of the estimator of the long-run variance converges to $\widehat{\Pi}_T(X)$, which itself converges to the long-run variance under suitable assumptions (e.g. Lu and Park (2014)). Thus, the estimator $\widehat{\Pi}_{T,\Delta}$ should stabilize as Δ becomes small. On the other hand, as Δ increases, we should see deviations of $\widehat{\Pi}_{T,\Delta}$ from $\widehat{\Pi}_T$ and hence also from the long-run variance itself. For what values of Δ do these deviations occur?

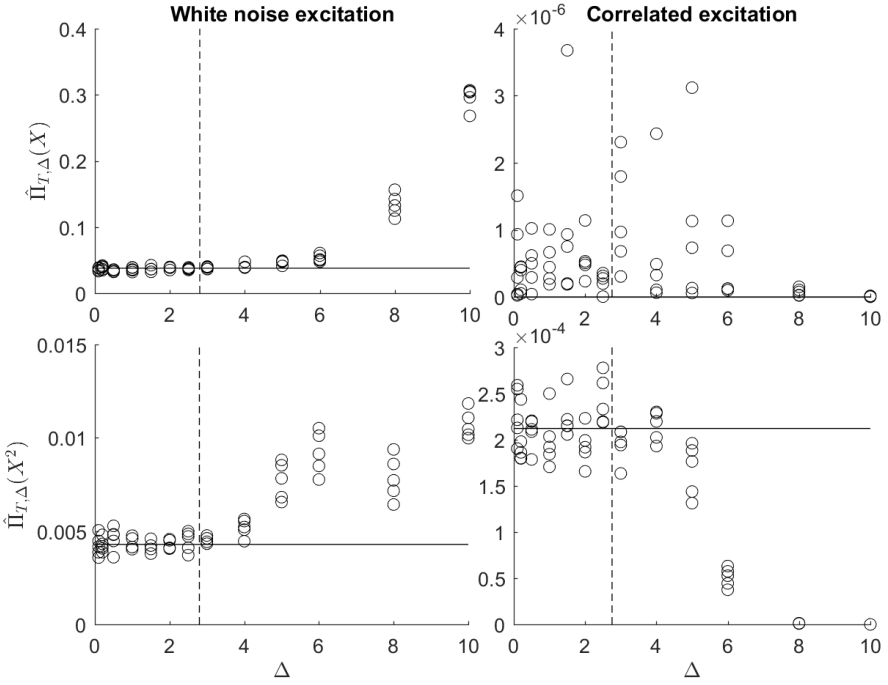

Figure 1: Estimated long-run variance $\widehat{\Pi}_{T,\Delta}$ against different choices of Δ. Five estimates are plotted for each Δ, using the data-driven approach with the QS kernel (see Section 2). Estimates are made for long-run variance of the original series $\widehat{\Pi}_{T,\Delta}(X)$ (top row), the squared series $\widehat{\Pi}_{T,\Delta}(X^2)$ (bottom row), under both white noise (left column) and correlated excitations (right column). Model parameters are as in Section 7.1, and the models are defined in Appendix A. Solid (horizontal) line gives true long-run variance; dashed (vertical) line marks $\frac{T_m}{4}$.

From numerical simulations across a range of oscillatory processes, we find that the estimator $\widehat{\Pi}_{T,\Delta}$ is quite stable up to about

$$\Delta_{\max} = \frac{T_m}{4}, \tag{6.28}$$

where T_m is the modal period of the oscillation associated with the frequency at which the spectrum is largest. This is illustrated in Figure 1 where the estimates $\widehat{\Pi}_{T,\Delta}(X)$ (top row) and $\widehat{\Pi}_{T,\Delta}(X^2)$ (bottom row) are plotted for several realizations of a linear oscillator with white noise excitation (left panel) and correlated excitation (right panel). The model parameters are the same as used in Section 7.1 below, and the models are defined in Appendix A. The value of $\Delta_{\max}$ in (6.28) is indicated by a vertical dashed line, and the true value of the long-run variance by a horizontal. (Note that the long-run variance $\Pi(X)$ is zero in the case of correlated excitation, in which case the corresponding plot only serves to show similar variability of the estimates up to $\Delta_{\max}$.)

Note also that $\Delta_{\max}$ in (6.28) is natural in the sense that $\int_{u_0}^{u_0+T_m} \Gamma_X(u)du$ (that is, the integral of $\Gamma_X(u)$ over its one approximate period of oscillation, as part of the long-run variance $\Pi(X)$) is expected to be approximated well enough by the integral discretized at step Δ, as long as $\Delta \leq \Delta_{\max}$. The latter is not meant as a rigorous statement. Note that if $\Delta > \Delta_{\max}$, e.g. $\Delta = T_m/2$ or T_m, then the discretization of the above integral could "pick up" only e.g. time points where $\Gamma_X(u)$

crosses zero (so that the discrete approximation will no longer be expected to be good).

6.5 Choice of time scale

Another practical issue, although not directly related to discretization, is the choice of a time scale. That is, whereas for discrete samples, the index scale is always the set of integers, a time scale for continuous-time process is subject to the practitioner's choice. For example, half an hour of data can be associated with $T = 1,800$ seconds (the time scale of seconds), as well as $T = 1/2$ hour (the time scale of hours). The practitioner should be aware of several implications of the choice of a time scale on the analysis.

We shall add a subscript, 1 or 2, to the quantities below to refer to the time scale 1 or 2, respectively (e.g. $T_1 = 1,800$ seconds and $T_2 = 1/2$ hours). The key observation here is that the value of the long-run variance $\Pi(X)$ in (1.2) actually depends on the chosen time scale. Indeed, note that

$$\Pi_2(X) = \int_{\mathbb{R}} \Gamma_{2,X}(u_2)du_2 = \frac{T_2}{T_1}\int_{\mathbb{R}} \Gamma_{2,X}\Big(\frac{T_2}{T_1}u_1\Big)du_1 = \frac{T_2}{T_1}\int_{\mathbb{R}} \Gamma_{1,X}(u_1)du_1 =: \frac{T_2}{T_1}\Pi_1(X), \qquad (6.29)$$

since the relation between the two time scales is $u_1 = (T_1/T_2)u_2$. But observe also that the time scale does not affect the variance of the sample mean, which would be used in the confidence intervals (see (2.8) and (2.9)), since

$$\frac{\Pi_2(X)}{T_2} = \frac{\Pi_1(X)}{T_1}.$$

The choice of a time scale affects similarly the estimators of the long-run variance. Indeed, note that, by arguing similarly as above,

$$\widehat{\Pi}_{2,T_2}(X) = \int_{-T_2}^{T_2} K\Big(\frac{u_2}{S_{T_2}}\Big)\Gamma_{T_2}(u_2)du_2 = \frac{T_2}{T_1}\int_{-T_1}^{T_1} K\Big(\frac{u_2}{S_{T_2}(T_1/T_2)}\Big)\Gamma_{T_1}(u_1)du_1 = \frac{T_2}{T_1}\widehat{\Pi}_{1,T_1}(X),$$

since $S_{T_2}(T_1/T_2) = S_{T_1}$ for optimal bandwidths $S_{T_1} = S_{\mathrm{opt},1,T_1}$ and $S_{T_2} = S_{\mathrm{opt},2,T_2}$. The latter relation follows by observing similarly that

$$C_{2,\nu}(X) = \frac{\Lambda_{2,\nu}(X)}{\Pi_2(X)} = \frac{(T_2/T_1)^{\nu+1}\Lambda_{1,\nu}(X)}{(T_2/T_1)\Pi_1(X)} = \Big(\frac{T_2}{T_1}\Big)^{\nu} C_{1,\nu}(X)$$

and hence indeed

$$S_{\mathrm{opt},2,T_2} = \Big(\frac{\nu K_\nu^2 C_{2,\nu}^2(X)}{\int K(x)^2dx}T_2\Big)^{1/(2\nu+1)} = \Big(\frac{\nu K_\nu^2 C_{1,\nu}^2(X)}{\int K(x)^2dx}\Big(\frac{T_2}{T_1}\Big)^{2\nu+1}T_1\Big)^{1/(2\nu+1)} = \frac{T_2}{T_1}S_{\mathrm{opt},1,T_1}.$$

Remark 6.2. The choice of a time scale does, however, affect one aspect of our analysis. The data-driven approach for estimating $\Pi(X)$ and $\Lambda_\nu(X)$ in the constant $C_\nu(X)$ in (2.5) (see the discussion following Remark 2.1) involves a preliminary choice of the bandwidth S_T, which was suggested as $\sqrt{T}$. Note that the latter choice depends on the time scale. Moreover, the choice of $\sqrt{T}$ was motivated by the fact that $\sqrt{T}$ should be smaller than T and that $\sqrt{T}/T \to 0$, as $T \to \infty$. A downside is that $\sqrt{T}$ is a meaningless choice in the case when $T < 1$ (e.g. $T = 1/2$ hour). As another possibility which does not have this problem and adapts naturally to the chosen time scale, a preliminary choice of S_T can be determined from the "decorrelation" method. This is an ad hoc method sometimes used in practice with S_T chosen as a time point where the sample autocorrelation function falls below a certain level (see Section 7.1 for a more detailed description).

7 Numerical results

In this section, we examine the methods proposed in Sections 2-6 through a simulation study (Section 7.1) and an application to real data (Section 7.2).

7.1 Simulation study

The simulation results presented in this section concern several synthetic processes often used in modeling oscillatory phenomena: a linear oscillator defined by (A.1) and (A.2), and a nonlinear (piecewise linear) oscillator defined by (A.1) and (A.3), in both cases with either a white noise excitation having a spectral density (A.4) or the correlated excitation having a spectral density (A.5). In the case of a linear oscillator, the parameters used are $\omega_0 = 0.6$, $\delta = 0.09$, and $\sigma = 0.07$ (the white noise excitation), $\sigma = 0.7$, $H_s = 9$, $T_1 = 11.595$ (the correlated excitation). In the case of a piecewise linear oscillator, we take $\omega_0 = 0.6, \delta = 0.09, k_1 = 0.1, x_m = \pi/6$, and the same parameters for the white noise and correlated excitations as in the linear case.

Tables 2–5 present simulation results for the linear and nonlinear oscillators with white noise and correlated excitations: Table 2 concerns the linear oscillator with a white noise excitation, Table 3 concerns the linear oscillator with a correlated excitation, Table 4 is for the piecewise linear oscillator with a white noise excitation, and Table 5 is for the piecewise linear oscillator with a correlated excitation. The first column in the tables indicates the length of the record (that is, $T = 50$ or 100 hours), with the associated simulated process sampled at $\Delta = 1/2$ second. The considered record lengths are typical to ship rolling applications. The second column indicates the kernel used in the estimation: QS for the Quadratic Spectral and B for the Bartlett kernel (see Section 2). The third column refers to the method used for estimation: "data" for the data-driven approach, "model" for the model-driven approach, "fixed" for estimation with a fixed bandwidth $S_T = \sqrt{T}$, "decor." for the decorrelation method and "self-norm" for the self-normalization approach. The decorrelation method is an ad hoc method sometimes used in practice where the bandwidth S_T is chosen as the cutoff point where an envelope of the autocorrelation function of the process crosses the level 0.05 (that is, 5% of the sample autocovariance) for the first time.

The next three columns in the tables present results when estimating the long-run variances $\Pi(X)$, $\Pi(X^2)$ and $\Pi(X, X^2)$. There are two entries in each box associated with a particular estimation scheme: the top entry gives the bias in estimation and the bottom entry gives the standard deviation in estimation, computed from 100 replications. We also note that the true value of $\Pi(X)$, $\Pi(X^2)$ and $\Pi(X, X^2)$ can be computed through explicit formulae in the case of a linear oscillator (see Appendix A), but that this is not the case for a piecewise linear oscillator, in which case we use an estimate from a record of 10,000 hours.

The fifth column of the tables presents the empirical coverage proportions of the proposed 95% confidence intervals for the mean $\mu(X)$ (the top entry in each box associated with a particular estimation scheme) and the standard deviation $\sigma(X)$ (the bottom entry in each box). As with the long-run variances above, the true standard deviation $\sigma(X)$ can be computed for the linear oscillator, but not for the nonlinear oscillator, in which case we use an estimate from a record of 10,000 hours. Finally, the last column of the tables gives the average half length of the corresponding confidence intervals (for the mean on the top, and for the standard deviation on the bottom). We also note that in Table 3, the last two rows report on the empirical coverages of the confidence intervals using quantiles, following Section 4.

Several conclusions can be drawn from Tables 2–5. First, the decorrelation method seems to

be the worst in general, both in terms of estimating long-run variances and coverages of confidence intervals. Second, the model-driven approach seems to perform best in general in terms of estimating long-run variances, in both linear and nonlinear cases. The performances of the data-driven and fixed approaches are difficult to discern, as are the confidence intervals among all 3 methods: model-driven, data-driven and fixed. Third, as expected, the confidence intervals for the mean $\mu(X)$ have 100% coverage in Table 3 – the appropriate shorter confidence intervals based on the quantile method have coverage close to 95%. Fourth, regarding the use of different kernels, estimation using the QS kernel seems generally superior to that for the Bartlett kernel, at least when the model-driven approach is used. Fifth, the self-normalization approach works quite well but as noted above, its confidence intervals are wider on average.

Tables 6–7 present simulation results in estimation of the long-run variance when multiple records are given. Table 6 is for a linear oscillator, and Table 7 is for a piecewise linear oscillator. The oscillator parameters are the same as those used in Tables 2–5. For both cases, $R = 10$ records of length $T_r = 5$ minutes each are considered. Three different methods of estimating $\Pi(X)$ (corresponding to X_t in the first column) and $\Pi(X^2)$ (X_t^2 in the first column) are examined: the proposed new ("mean-all"), the average ("separate") and the direct methods (see Section 5). The Kernel and Method columns are the same as in Tables 2–5, except that the decorrelation method is excluded. The entries in each box associated with a particular estimation scheme now indicate the bias (top entry), the standard deviation (middle entry), and the mean-squared error (bottom entry). The entries with the smallest mean-squared errors are indicated in bold.

It can be seen from Tables 6–7 that the proposed estimator (mean-all) has superior performance in the largest number of cases and always performs better when the model-based approach and the QS kernel are used. This approach/kernel was suggested above as superior for single records.

We have also produced tables in the case of multiple records that are analogous to Tables 2–5 but will not include them for the shortness sake. Few observations, however, stood out from these unreported simulations. First, the negative bias in estimation of the long-run variance (as in several cases in Tables 6–7) led naturally to smaller coverage proportions. Second, the self-normalization approach again produced wider confidence intervals on average but the increase factor over the long-run variance approach was larger for multiple records than in the case of a single record.

7.2 Data application

We illustrate here the proposed methodology on the data generated by a high-fidelity ship motion simulation code (more specifically, Large Amplitude Motion Program or LAMP of Lin and Yue (1991)). The data in question concerns loads at a particular point of a ship. The time plot of the data is depicted in Figure 2 (left). The duration of the record is $T = 819$ seconds (with the first 20 seconds discarded), and we consider a sampling rate ranging from $\Delta = 0.02$ to $\Delta = 0.16$ seconds. We are interested in providing confidence intervals for the mean and the standard deviation of the underlying process.

As a first step, we need to decide whether the data points to the degenerate case. As discussed in Section 4, the degenerate case is associated with the fact that the process $\int_0^t (X_s - \mu(X))ds$ is stationary (and vice-versa, the non-degenerate case with the process being non-stationary). The time plot of the sample analogue of the process, namely, the discrete version of $\int_0^t (X_s - \overline{X}_T)ds$, is given in Figure 2 (right). To see whether the non-degenerate case can be assumed, we postulate it as a null hypothesis to be tested. We use the well-known augmented Dickey-Fuller (ADF) test for the null hypothesis – see e.g. Chapter 3 in Pfaff (2008). The test statistic value is -24.3202, and

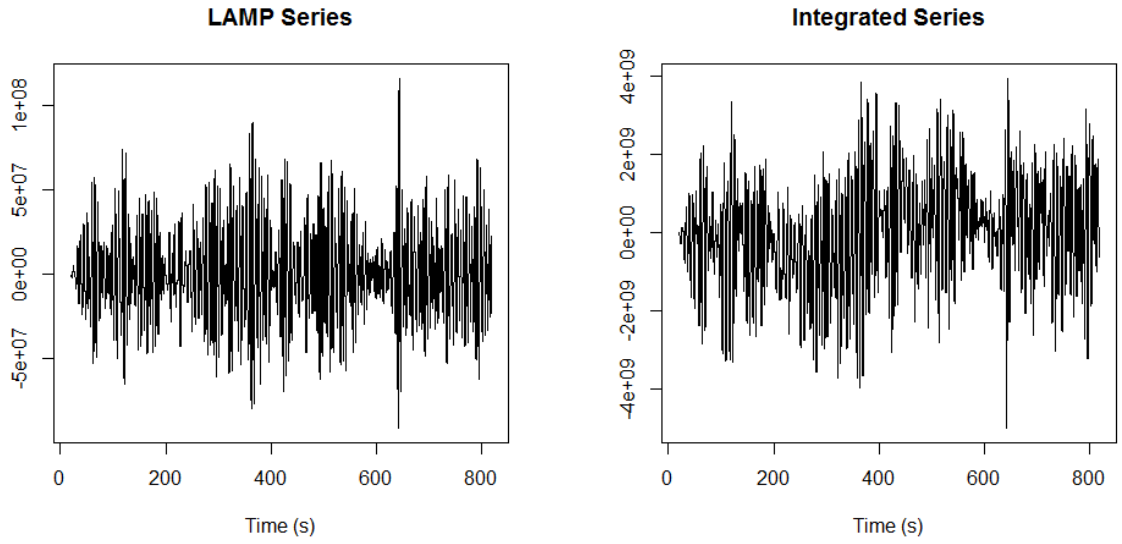

Figure 2: Time series data simulated from LAMP. Left: the series itself with $\Delta = 0.02$. Right: the discrete version of $\int_0^t (X_s - \overline{X}_T)ds$.

the critical value is -2.58 at $\alpha = 1\%$ (-1.95 at $\alpha = 5\%$). Based on these values, we reject the null hypothesis, i.e. conclude that the data is consistent with the degenerate case.

Proceeding with the methods proposed in the degenerate case, the sample mean of the process with $\Delta = 0.02$ (the smallest available) is 313739.6, and the corresponding 95% confidence intervals for the mean for several choices of Δ are shown in Figure 3 (left).

To construct confidence intervals for the standard deviation for the LAMP series, we need only look at the squared series to estimate its long-run variance (see Section 4). The squared series and its integrated series are plotted in Figure 4. The ADF test gives a test statistic value of -2.3433; as expected, we consider the squared series to be non-degenerate and proceed with the methods proposed in Section 3, using the QS kernel and a fixed bandwidth of $\sqrt{T}$. The sample standard deviation of the process with $\Delta = 0.02$ is 28119779, and the corresponding 95% confidence intervals are shown in Figure 3 (right). The confidence intervals for the standard deviation obtained from the self-normalization approach were, in fact, similar to those in Figure 3 and will not be reported for the shortness sake.

8 Conclusions

The focus of this work has been on inference (i.e. confidence intervals) for the mean and variance of random oscillatory processes by making use of either estimation of the long-run variance of the system or the self-normalization approach. We considered processes with both positive long-run variance, as well as the "degenerate case" with zero long-run variance; analysis of the latter required different estimation techniques. Additionally, we examined the case of multiple independent records of the same stochastic process, from which we developed a more promising method of estimation. We presented numerical results of the proposed methods through a simulation study, demonstrating

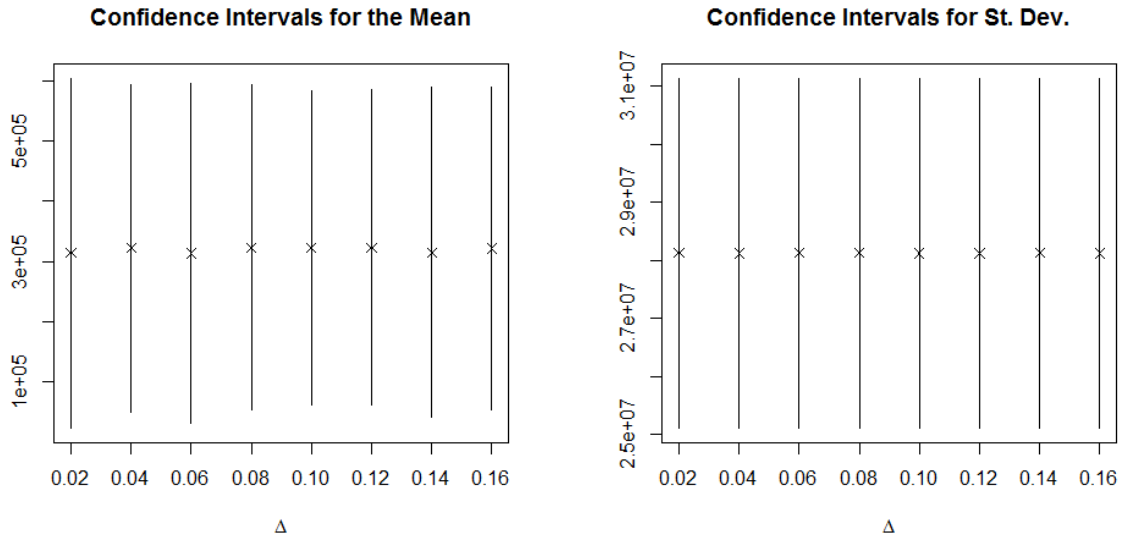

Figure 3: Confidence intervals for the mean from the LAMP series for different sampling rates. Left: confidence intervals for the mean. Right: confidence intervals for the standard deviation.

the performance of these estimators.

Acknowledgments

The work described in this paper has been funded by the Office of Naval Research (ONR) under Dr. Thomas Fu and Dr. Woei-Min Lin and by the Naval Surface Warfare Center Carderock Division (NSWCCD) Independent Applied Research (IAR) program. Participation of Prof. Pipiras was facilitated by the Summer Faculty Program supported by ONR and managed by NSWCCD under Dr. Jack Price, who also manages IAR program. Participation of Mr. Glozter was facilitated by the NSWCCD Naval Research Enterprise Internship Program (NREIP) program managed by Ms. Rachel Luu. Prof. Pipiras also acknowledges the partial support of the NSF grant DMS-1712966 at the University of North Carolina.

A Oscillatory systems of interest

We describe here several oscillatory processes that are used and referred to in this work. In the simulations of Section 7.1, we consider an *oscillator* X_t satisfying the general equation

$$\ddot{X}_t + 2\delta\dot{X}_t + r(X_t) = Z_t. \tag{A.1}$$

Here, $\delta > 0$ is a damping parameter, $r(x)$ is a restoring force and Z_t is an external excitation. A *linear oscillator* corresponds to a linear restoring function

$$r(x) = w_0^2 x, \tag{A.2}$$

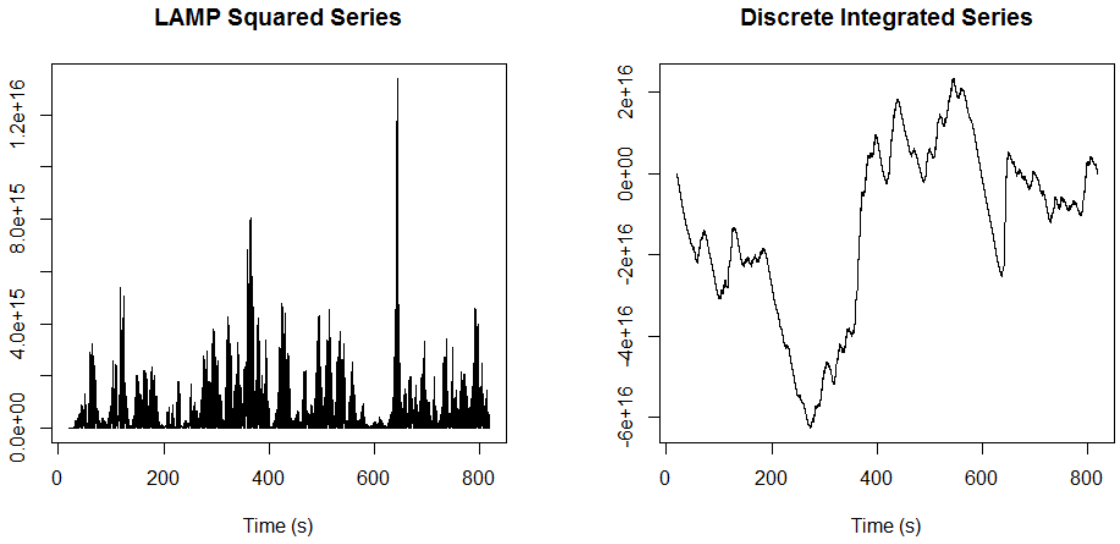

Figure 4: Time series data simulated from LAMP and squared. Left: the squared series itself with $\Delta = 0.02$. Right: the discrete version of $\int_0^t (X_s^2 - \overline{X_T^2})ds$.

where w_0 is a natural frequency parameter. A *nonlinear oscillator* is associated with a nonlinear restoring force. In Section 7.1, we use a *piecewise linear oscillator* with a restoring force

$$r(x) = \begin{cases} -k_1 w_0^2 (x + x_m) - w_0^2 x_m, & x < -x_m, \\ w_0^2 x, & -x_m < x < x_m, \\ -k_1 w_0^2 (x - x_m) + w_0^2 x_m, & x > x_m, \end{cases} \tag{A.3}$$

that is, where the restoring force has a negative slope ($-k_1 w_0^2$) in the nonlinear regime $|x| > x_m$. Though other nonlinear oscillators (e.g. the Duffing oscillator) could be considered as well.

Two forms of the external excitation are considered. First, there is a white noise excitation $Z_t = \sigma \dot{W}_t$, where W_t is a standard Brownian motion and $\sigma > 0$ is the parameter controlling the strength of the excitation. Its spectral density can be thought as

$$S_Z(w) = \sigma^2. \tag{A.4}$$

Second, motivated by ship rolling applications, we also consider a stationary Gaussian excitation having a spectral density

$$S_Z(w) = w_0^4 \Big(\frac{w^2}{g}\Big)^2 \frac{A}{w^5} e^{-\frac{B}{w^4}}, \quad w > 0, \tag{A.5}$$

where w_0 is as in (A.2) and (A.3) and $g = 9.807$ is the gravitational acceleration. The parameters A and B are taken as $A = 173 H_s^2 T_1^{-4}$ and $B = 691 T_1^{-4}$, where H_s is a significant wave height, i.e. twice the amplitude in meters of the highest one-third of waves, and T_1 is the period corresponding to the mean frequency of waves.

We also note that the spectral density of the linear oscillator with excitation Z is given by

$$S_X(w) = \frac{S_Z(w)}{(w_0^2 - w^2)^2 + (2\delta w)^2}. \tag{A.6}$$

Since $S_Z(0) = 0$ for the correlated excitation with the spectral density (A.5), we have $S_X(0) = 0$ for the linear oscillator with the correlated excitation. Thus, in view of (4.1), this oscillator falls into the degenerate case.

A convenient fact about a linear oscillator with a white noise excitation is that its autocovariance function can be computed explicitly as follows. In view of (2.7) and (A.6), we have

$$\Gamma_X(h) = \frac{1}{2\pi}\int_{\mathbb{R}} \cos(hw) S_X(w) dw = \frac{\sigma^2}{\pi}\int_0^\infty \frac{\cos(hw)}{(w_0^2 - w^2)^2 + (2\delta w)^2} dw$$

$$= \frac{\sigma^2}{\pi}\int_0^\infty \frac{\cos(hw)}{w^4 + 2w_0^2 w^2 (2(\delta/w_0)^2 - 1) + w_0^4} dw = \frac{\sigma^2}{\pi}\int_0^\infty \frac{\cos(hw)}{w^4 + 2w_0^2 w^2 \cos(2t) + w_0^4} dw,$$

where $\cos(2t) = 2(\delta/w_0)^2 - 1$ with $0 < t < \pi/2$, assuming that

$$\delta < w_0.$$

Then, by using Formula 3.733.1 in Gradshteyn and Ryzhik (2014), p. 428,

$$\Gamma_X(h) = \frac{\sigma^2}{2w_0^3 \sin(2t)} e^{-hw_0 \cos(t)} \sin(t + hw_0 \sin(t))$$

$$= \frac{\sigma^2}{4w_0 \delta \sqrt{w_0^2 - \delta^2}} e^{-h\delta} \sin(t + h\sqrt{w_0^2 - \delta^2}), \tag{A.7}$$

by using the facts that $w_0 \cos(t) = \delta$ and $w_0 \sin(t) = \sqrt{w_0^2 - \delta^2}$. This also allows one to compute explicitly the long-run variance of the process as: by using Formula 3.893.1 in Gradshteyn and Ryzhik (2014), p. 486,

$$\Pi(X) = \int_0^\infty \Gamma_X(h) dh = \frac{\sigma^2}{w_0^3 \sin(2t)} \int_0^\infty e^{-hw_0 \cos(t)} \sin(t + hw_0 \sin(t)) dh$$

$$= \frac{\sigma^2}{w_0^3 \sin(2t)} \cdot \frac{1}{w_0^2 \cos^2(t) + w_0^2 \sin^2(t)} \Big(w_0 \sin(t) \cos(t) + w_0 \cos(t) \sin(t) \Big) = \frac{\sigma^2}{w_0^4}. \tag{A.8}$$

Note that this is also the same as $S_X(0)$. One can also compute $\Lambda_2(X)$ in (2.5)–(2.6) explicitly (see Remark 2.1) as

$$\Lambda_2(X) = (-1) \frac{d^2 S_X(\omega)}{d\omega^2}\Big|_{w=0} = \frac{4\sigma^2(2\delta^2 - \omega_0^2)}{\omega_0^8}. \tag{A.9}$$

In the Gaussian case, which we assume for a linear oscillator with a white or uncorrelated excitation, long-run variances associated with the squares of the process can be calculated by using the well-known relations

$$\Gamma_{X^2}(h) = 2(\Gamma_X(h))^2, \quad \Gamma_{X,X^2}(h) = 0. \tag{A.10}$$

For nonlinear oscillators (with any excitation), no explicit formulae are known for either the spectral density or the autocovariance function.

References

D. W. K. Andrews. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):817–858, 1991.

H. Fan, T. Soderstrom, M. Mossberg, B. Carlsson, and Y. Zou. Estimation of continuous-time AR process parameters from discrete-time data. *IEEE Transactions on Signal Processing*, 47(5): 1232–1244, 1999.

M. Gitterman. *The Noisy Oscillator: The First Hundred Years, From Einstein Until Now*. World Scientific, 2005.

I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic press, 2014.

H. Kirshner, M. Unser, and J. P. Ward. On the unique identification of continuous-time autoregressive models from sampled data. *IEEE Transactions on Signal Processing*, 62(6):1361–1376, 2014.

J. Lee. Long-run variance estimation for linear processes under possible degeneracy. *Journal of Economic Theory and Econometrics*, 21(1):1–22, 2010.

Z. Liang and G. C. Lee. *Random Vibration: Mechanical, Structural, and Earthquake Engineering Applications*. CRC Press/Taylor & Francis Group, 2015.

N. Lin and S. V. Lototsky. Undamped harmonic oscillator driven by additive Gaussian white noise: a statistical analysis. *Communications on Stochastic Analysis*, 5(1):233–250, 2011.

N. Lin and S. V. Lototsky. Second-order continuous-time non-stationary Gaussian autoregression. *Statistical Inference for Stochastic Processes*, 17(1):19–49, 2014.

W. M. Lin and D. K. P. Yue. Numerical solutions for large amplitude ship motions in the time-domain. In *Proceedings of the 18th Symposium on Naval Hydrodynamics*, pages 41–66, Ann Arbor, 1991.

I. N. Lobato. Testing that a dependent process is uncorrelated. *Journal of the American Statistical Association*, 96(455):1066–1076, 2001.

Y. Lu and J. Y. Park. Estimation of longrun variance of continuous time stochastic process using discrete sample. Preprint, 2014.

J. I. Neimark and P. S. Landa. *Stochastic and Chaotic Oscillations*, volume 77. Springer Science & Business Media, 2012.

E. Parzen. On consistent estimates of the spectrum of a stationary time series. *Annals of Mathematical Statistics*, 28:329–348, 1957.

B. Pfaff. *Analysis of Integrated and Cointegrated Time Series with R*. Springer Science & Business Media, 2008.

D. T. Pham. Estimation of continuous-time autoregressive model from finely sampled data. *IEEE Transactions on Signal Processing*, 48(9):2576–2584, 2000.

X. Shao. Self-normalization for time series: a review of recent developments. *Journal of the American Statistical Association*, 110(512):1797–1817, 2015.

T. Soderstrom, H. Fan, B. Carlsson, and S. Bigi. Least squares parameter estimation of continuous-time ARX models from discrete-time data. *IEEE Transactions on Automatic Control*, 42(5): 659–673, 1997.

Hours	Kernel	Method	$\Pi(X)$	$\Pi(X^2)$	$\Pi(X, X^2)$	Coverage prop.	Interval length
50	QS	data	7.27e-02	1.29e+00	2.38e-01	0.96	1.29e-02
			4.97e-01	1.28e+01	1.88e+00	0.91	1.44e-02
50	QS	model	7.74e-02	-3.41e+00	6.66e-02	0.95	1.29e-02
			1.48e-01	5.45e+00	6.01e-01	0.93	1.42e-02
50	QS	fixed	-7.09e-02	-2.49e+00	-1.84e-01	0.95	1.28e-02
			5.13e-01	1.20e+01	2.03e+00	0.96	1.42e-02
50	QS	decor.	2.90e-01	-1.10e+01	3.48e-02	0.98	1.31e-02
			1.02e-01	4.56e+00	3.84e-01	0.92	1.39e-02
50	B	data	4.96e-01	-1.26e+01	1.49e-01	0.97	1.32e-02
			5.93e-01	1.17e+01	6.00e-01	0.95	1.38e-02
50	B	model	1.89e-01	-5.62e+00	-5.67e-03	0.96	1.30e-02
			1.99e-01	5.85e+00	7.69e-01	0.95	1.41e-02
50	B	fixed	6.59e-02	-2.31e+00	-9.29e-02	0.98	1.29e-02
			4.08e-01	1.06e+01	1.59e+00	0.95	1.42e-02
50	B	decor.	1.00e+00	-2.50e+01	-6.63e-03	0.98	1.36e-02
			1.18e-01	3.86e+00	3.70e-01	0.94	1.33e-02
50	-	self-norm	-	-	-	0.95	1.72e-02
						0.97	2.05e-02
100	QS	data	7.21e-02	4.34e-01	-1.69e-01	0.93	9.11e-03
			4.40e-01	9.59e+00	1.67e+00	0.93	1.02e-02
100	QS	model	7.10e-02	-2.00e+00	-1.63e-02	0.95	9.12e-03
			1.14e-01	4.04e+00	4.23e-01	0.93	1.01e-02
100	QS	fixed	1.71e-02	-7.73e-02	2.27e-02	0.98	9.08e-03
			3.78e-01	9.48e+00	1.69e+00	0.95	1.01e-02
100	QS	decor.	2.91e-01	-1.09e+01	-5.00e-03	0.96	9.24e-03
			8.37e-02	3.02e+00	2.69e-01	0.93	9.83e-03
100	B	data	3.56e-01	-8.54e+00	2.57e-02	0.96	9.28e-03
			3.44e-01	8.46e+00	4.74e-01	0.98	9.89e-03
100	B	model	1.11e-01	-3.11e+00	5.07e-02	1.00	9.14e-03
			2.07e-01	4.87e+00	5.78e-01	0.96	1.01e-02
100	B	fixed	-2.52e-02	-2.40e+00	6.85e-02	0.97	9.06e-03
			3.58e-01	7.88e+00	1.13e+00	0.96	1.01e-02
100	B	decor.	9.99e-01	-2.60e+01	5.46e-03	0.99	9.64e-03
			7.41e-02	2.33e+00	2.57e-01	0.94	9.38e-03
100	-	self-norm	-	-	-	0.95	1.13e-02
						0.98	1.440e-02

Table 2: Simulation results for the linear oscillator with white noise excitation. See Section 7.1 for discussion.

Hours	Kernel	Method	$\Pi(X)$	$\Pi(X^2)$	$\Pi(X, X^2)$	Coverage prop.	Interval length
50	QS	data	1.47e-07	1.07e-04	-2.25e-08	1.00	1.57e-06
			1.55e-07	1.72e-05	1.97e-07	0.95	5.38e-04
50	QS	model	1.58e-07	9.83e-05	-4.84e-09	1.00	1.63e-06
			1.42e-07	6.52e-06	8.54e-08	0.96	5.27e-04
50	QS	fixed	1.27e-07	1.10e-04	-1.59e-09	1.00	1.44e-06
			1.26e-07	1.62e-05	1.88e-07	0.94	5.42e-04
50	QS	decor.	1.23e-07	8.73e-05	-1.44e-09	1.00	1.43e-06
			1.24e-07	5.57e-06	4.63e-08	0.95	5.12e-04
50	B	data	5.53e-05	1.04e-04	2.20e-08	1.00	3.43e-05
			2.26e-06	1.40e-05	1.71e-07	0.96	5.35e-04
50	B	model	1.14e-04	1.02e-04	-4.13e-10	1.00	4.94e-05
			1.02e-06	1.10e-05	9.69e-08	0.93	5.31e-04
50	B	fixed	5.72e-05	1.03e-04	1.47e-08	1.00	3.49e-05
			6.44e-07	1.38e-05	1.10e-07	0.95	5.33e-04
50	B	decor.	1.01e-03	6.35e-05	-1.14e-08	1.00	1.47e-04
			3.08e-05	4.52e-06	2.87e-07	0.95	4.80e-04
50	-	self-norm	-	-	-	1.00	5.38e-06
						0.95	6.90e-04
50	-	quantile	-	-	-	0.943	1.70e-08
100	QS	data	7.61e-08	1.09e-04	1.06e-08	1.00	8.29e-07
			6.20e-08	1.19e-05	1.14e-07	0.88	3.82e-04
100	QS	model	6.92e-08	1.00e-04	-1.76e-09	1.00	7.86e-07
			5.23e-08	4.98e-06	2.84e-08	0.89	3.74e-04
100	QS	fixed	6.79e-08	1.07e-04	-1.96e-08	1.00	7.59e-07
			6.99e-08	1.32e-05	1.05e-07	0.90	3.81e-04
100	QS	decor.	6.15e-08	8.64e-05	4.90e-10	1.00	7.12e-07
			6.22e-08	4.60e-06	1.98e-08	0.88	3.62e-04
100	B	data	4.34e-05	1.06e-04	-9.07e-09	1.00	2.15e-05
			1.35e-06	1.01e-05	8.62e-08	0.96	3.80e-04
100	B	model	9.05e-05	1.04e-04	2.29e-10	1.00	3.11e-05
			5.86e-07	8.02e-06	7.00e-08	0.98	3.78e-04
100	B	fixed	4.05e-05	1.05e-04	-5.15e-10	1.00	2.08e-05
			2.90e-07	9.86e-06	1.29e-07	0.93	3.78e-04
100	B	decor.	1.01e-03	6.31e-05	9.26e-09	1.00	1.04e-04
			2.15e-05	4.00e-06	2.23e-07	0.91	3.39e-04
100	-	self-norm	-	-	-	1.00	2.53e-06
						0.92	5.18e-04
100	-	quantile	-	-	-	0.955	8.47e-09

Table 3: Simulation results for the linear oscillator with correlated excitation. See Section 7.1 for discussion.

Hours	Kernel	Method	$\Pi(X)$	$\Pi(X^2)$	$\Pi(X, X^2)$	Coverage prop.	Interval length
50	QS	data	2.84e-05	2.15e-03	-1.22e-04	0.92	9.29e-04
			2.76e-03	3.41e-03	1.73e-03	0.95	2.09e-03
50	QS	model	5.01e-04	-5.76e-04	-4.41e-06	0.97	9.35e-04
			8.56e-04	1.73e-03	8.94e-04	0.96	2.02e-03
50	QS	fixed	3.43e-06	2.52e-03	4.08e-04	0.99	9.29e-04
			2.61e-03	3.52e-03	1.99e-03	0.94	2.10e-03
50	QS	decor.	5.39e-04	-1.64e-04	6.31e-05	0.91	9.36e-04
			8.91e-04	1.76e-03	7.66e-04	0.94	2.03e-03
50	B	data	4.17e-03	-1.49e-03	-1.53e-04	0.97	9.76e-04
			2.73e-03	3.04e-03	8.73e-04	0.94	1.99e-03
50	B	model	1.84e-03	1.22e-04	-8.53e-05	0.94	9.50e-04
			1.42e-03	2.47e-03	1.22e-03	0.95	2.04e-03
50	B	fixed	7.85e-04	2.07e-03	2.21e-05	0.93	9.38e-04
			2.29e-03	3.00e-03	1.62e-03	0.97	2.08e-03
50	B	decor.	6.52e-03	-4.16e-03	-1.44e-05	0.98	1.00e-03
			8.30e-04	1.61e-03	7.75e-04	0.95	1.92e-03
50	-	self-norm	-	-	-	0.96	1.21e-03
						0.96	2.72e-03
100	QS	data	-1.22e-04	2.79e-03	3.64e-04	0.93	6.56e-04
			2.54e-03	2.71e-03	1.79e-03	0.96	1.49e-03
100	QS	model	4.45e-04	1.57e-04	-5.03e-05	1.00	6.61e-04
			7.17e-04	1.39e-03	5.33e-04	0.98	1.44e-03
100	QS	fixed	1.85e-04	2.14e-03	-9.14e-05	0.88	6.59e-04
			2.01e-03	3.03e-03	1.67e-03	0.94	1.48e-03
100	QS	decor.	5.40e-04	-4.02e-04	-6.01e-05	0.97	6.62e-04
			6.51e-04	1.43e-03	5.87e-04	0.94	1.43e-03
100	B	data	2.78e-03	-3.98e-04	6.14e-05	0.95	6.79e-04
			1.79e-03	2.12e-03	6.34e-04	0.96	1.43e-03
100	B	model	9.84e-04	1.57e-03	-4.11e-05	0.97	6.65e-04
			1.19e-03	1.65e-03	7.90e-04	0.94	1.47e-03
100	B	fixed	5.89e-04	2.63e-03	1.95e-04	0.96	6.62e-04
			1.78e-03	2.19e-03	1.49e-03	0.96	1.48e-03
100	B	decor.	6.71e-03	-4.33e-03	-4.11e-05	0.95	7.10e-04
			5.82e-04	1.23e-03	4.92e-04	0.94	1.35e-03
100	-	self-norm	-	-	-	0.96	8.52e-04
						0.97	1.93e-03

Table 4: Simulation results for the piecewise linear oscillator with white noise excitation. See Section 7.1 for discussion.

Hours	Kernel	Method	$\Pi(X)$	$\Pi(X^2)$	$\Pi(X, X^2)$	Coverage prop.	Interval length
50	QS	data	-1.43e-03	2.81e-03	1.07e-05	0.94	5.35e-05
			1.59e-05	4.17e-03	1.63e-04	0.93	2.38e-03
50	QS	model	-1.43e-03	-1.64e-03	-1.49e-05	0.98	5.51e-05
			1.27e-05	2.22e-03	1.17e-04	0.93	2.28e-03
50	QS	fixed	-1.43e-03	3.61e-03	1.50e-05	0.91	5.36e-05
			1.64e-05	4.42e-03	1.93e-04	0.90	2.40e-03
50	QS	decor.	-1.43e-03	-8.01e-04	-4.33e-06	0.96	5.49e-05
			1.42e-05	2.15e-03	1.26e-04	0.91	2.30e-03
50	B	data	2.03e-03	-1.19e-03	1.49e-07	1.00	2.75e-04
			8.72e-04	2.69e-03	1.13e-04	0.98	2.28e-03
50	B	model	-3.22e-04	1.14e-03	-3.63e-07	1.00	1.63e-04
			2.26e-05	2.91e-03	1.37e-04	0.98	2.34e-03
50	B	fixed	-7.23e-04	1.83e-03	2.60e-05	1.00	1.34e-04
			2.16e-05	3.58e-03	1.39e-04	0.99	2.36e-03
50	B	decor.	5.29e-03	-6.35e-03	-5.88e-06	1.00	3.82e-04
			2.84e-04	1.99e-03	1.09e-04	0.91	2.16e-03
50	-	self-norm	-	-	-	1.00	1.66e-06
						0.90	3.05e-04
100	QS	data	-1.43e-03	2.61e-03	1.18e-05	0.94	3.79e-05
			1.16e-05	3.35e-03	1.44e-04	0.95	1.68e-03
100	QS	model	-1.43e-03	-6.35e-04	2.70e-06	0.97	3.89e-05
			9.01e-06	1.43e-03	7.49e-05	0.97	1.63e-03
100	QS	fixed	-1.43e-03	2.19e-03	2.59e-05	0.95	3.79e-05
			1.24e-05	3.74e-03	1.19e-04	0.93	1.67e-03
100	QS	decor.	-1.43e-03	-7.20e-04	2.40e-05	0.96	3.89e-05
			9.19e-06	1.42e-03	8.64e-05	0.91	1.63e-03
100	B	data	1.37e-03	-8.13e-04	-1.46e-06	1.00	1.76e-04
			7.78e-04	1.89e-03	7.93e-05	0.96	1.62e-03
100	B	model	-5.51e-04	1.62e-03	-9.09e-06	1.00	1.04e-04
			1.44e-05	2.35e-03	9.87e-05	0.94	1.66e-03
100	B	fixed	-9.26e-04	2.44e-03	-3.13e-06	1.00	8.27e-05
			1.46e-05	3.07e-03	1.22e-04	0.90	1.68e-03
100	B	decor.	5.27e-03	-6.09e-03	-1.43e-06	1.00	2.70e-04
			1.78e-04	1.26e-03	8.96e-05	0.94	1.53e-03
100	-	self-norm	-	-	-	1.00	8.35e-07
						0.90	2.17e-04

Table 5: Simulation results for the piecewise linear oscillator with correlated excitation. See Section 7.1 for discussion.

Series	Kernel	Method	Separate	Mean-all	Direct
X_t	QS	data	-0.6492	0.0002	1.4624
			0.0410	0.0614	11.1956
			0.4624	**0.0614**	13.3342
X_t	QS	model	-0.7979	-0.2130	2.0293
			0.0087	0.0125	8.9627
			0.6453	**0.0578**	13.0807
X_t	QS	fixed	-0.0885	0.0168	1.6874
			0.0746	0.1122	7.5065
			0.0825	0.1125	10.3537
X_t	B	data	-0.7729	-1.2907	1.9251
			0.3506	0.2105	9.5462
			0.9480	1.8765	13.2522
X_t	B	model	-0.0540	-0.4557	1.5672
			0.0291	0.0191	9.3266
			0.0320	0.2267	11.7828
X_t	B	fixed	0.0735	-0.1111	1.7914
			0.0733	0.0494	8.7393
			0.0787	**0.0618**	11.9486
X_t^2	QS	data	1.8409	0.7258	3.4816
			45.1346	44.7539	8442.4808
			48.5234	**45.2807**	8454.6025
X_t^2	QS	model	11.4759	11.2818	9.0762
			18.6724	18.7036	6326.1594
			150.3685	**145.9823**	6408.5360
X_t^2	QS	fixed	3.4290	1.8819	-6.6153
			66.8617	67.4763	8543.4466
			78.6199	**71.0178**	8587.2089
X_t^2	B	data	33.3687	32.0303	-10.4373
			58.4451	56.7885	6685.3033
			1171.9130	**1082.7290**	6794.2412
X_t^2	B	model	13.8039	13.4828	10.6049
			27.7812	27.8504	5690.6623
			218.3299	**209.6351**	5803.1269
X_t^2	B	fixed	4.5421	3.2672	-12.0358
			44.0204	44.0750	6730.3875
			64.6511	**54.7495**	6875.2490

Table 6: Simulation results for multiple records of the linear oscillator with white noise excitation. See Section 7.1 for discussion.

Series	Kernel	Method	Separate	Mean-all	Direct
X_t	QS	data	-8.9125e-03	-4.8590e-04	-1.7689e-04
			7.4359e-06	3.3759e-06	3.4772e-04
			8.6869e-05	**3.6120e-06**	3.4775e-04
X_t	QS	model	-1.0305e-02	-1.1530e-03	7.1772e-04
			3.6025e-07	4.5355e-07	2.5875e-04
			1.0655e-04	**1.7830e-06**	2.5927e-04
X_t	QS	fixed	-1.6811e-03	-5.9613e-05	-7.3207e-04
			1.3867e-06	2.1663e-06	3.2782e-04
			4.2126e-06	**2.1699e-06**	3.2835e-04
X_t	B	data	-6.0523e-03	-1.3383e-02	2.6429e-04
			3.1945e-05	2.9260e-05	3.5582e-04
			6.8575e-05	2.0835e-04	3.5589e-04
X_t	B	model	-1.7149e-04	-3.5594e-03	-3.1205e-03
			1.1444e-06	7.8580e-07	4.0284e-04
			1.1738e-06	1.3455e-05	4.1258e-04
X_t	B	fixed	-4.9586e-05	-2.2208e-03	6.2363e-04
			2.2776e-06	1.4943e-06	2.8804e-04
			2.2800e-06	6.4264e-06	2.8843e-04
X_t^2	QS	data	-1.4019e-03	-1.6416e-03	-4.0445e-04
			4.4542e-06	4.5681e-06	4.6088e-04
			6.4195e-06	7.2628e-06	4.6104e-04
X_t^2	QS	model	3.4570e-03	3.3914e-03	-3.9850e-03
			2.1546e-06	2.1500e-06	2.6057e-04
			1.4106e-05	**1.3652e-05**	2.7645e-04
X_t^2	QS	fixed	-1.6560e-03	-2.0158e-03	-5.6502e-03
			6.1992e-06	6.4010e-06	3.9172e-04
			8.9414e-06	1.0464e-05	4.2365e-04
X_t^2	B	data	8.1631e-03	7.7285e-03	-7.1866e-03
			6.3211e-06	6.3417e-06	4.3724e-04
			7.2957e-05	**6.6071e-05**	4.8889e-04
X_t^2	B	model	3.4164e-03	3.3269e-03	1.6838e-03
			3.1475e-06	3.2081e-06	3.4893e-04
			1.4819e-05	**1.4276e-05**	3.5176e-04
X_t^2	B	fixed	-4.1757e-04	-6.8928e-04	-3.2341e-03
			4.0635e-06	4.1038e-06	3.3397e-04
			4.2379e-06	4.5789e-06	3.4443e-04

Table 7: Simulation results for multiple records of the piecewise linear oscillator with correlated excitation. See Section 7.1 for discussion.