The Quality of Lagged Products and Autoregressive Yule–Walker Models as Autocorrelation Estimates

Piet M. T. Broersen

***Abstract*—The sample autocorrelation function is defined by the mean lagged products (LPs) of random observations. It is the inverse Fourier transform of the raw periodogram. Both contain the same information, and the quality of the full-length sample autocorrelation to represent random data is as poor as that of a raw periodogram. The autoregressive (AR) Yule–Walker method uses LP autocorrelation estimates to compute AR parameters as a parametric model for the autocorrelation. The order of the AR model can be taken as the full LP length, or it can be determined with an order selection criterion. However, the autocorrelation function can more accurately be estimated with a general parametric time-series method. This parametric estimate of the autocorrelation function always has better accuracy than the LP estimates. The LP autocorrelation function is as long as the observation window, but parametric estimates will eventually die out. They allow an objective answer to the question of how long the autocorrelation function really is.**

***Index Terms*—Autoregressive (AR) process, correlation, identification, order selection, spectral estimation, time-series model.**

I. Introduction

THE PROCESSING of random signals became a useful computational tool after the rediscovery of the fast Fourier transform (FFT) algorithm [1]. The square of the absolute value of the FFT of the signal is the periodogram [2]. This algorithm enabled the routine Fourier analysis of large sets of stochastic data. The inverse Fourier transform of the periodogram is the lagged-product (LP) autocovariance estimate. It is computationally efficient to calculate the autocorrelation function as the inverse transform of the periodogram. Like the periodogram for spectra, it has been treated in the literature as if it is the obvious estimator for the autocorrelation function for random data [2]–[6].

The length of the estimated LP autocorrelation is determined either by the sample size or by the choice of the data analyst. It is not determined by the character of the data. This may cause problems in the analysis of the correlation length of measurement data or the lifetime of physical processes. Recently, it has been discovered that the lifetime of turbulent phenomena is sometimes finite [7].

Only *relative* spectral differences or squared *logarithmic* differences are advisable as measures for the spectral accuracy [8]. Unfortunately, the sum of squared differences between true and estimated spectra or autocorrelations is certainly not an adequate measure [9], although its use seems to be popular. The accuracy in the time domain can be defined as a squared prediction error (PE) on the level of the data, as a cepstral distance for autocorrelations, or as a sum of squared differences of the logarithms for spectral estimates. These three measures and several other similar useful measures are strongly related [9], [10]. They can be computed in simulation studies if the true process characteristics are known.

Parametric time-series models are now the preferred methods for spectral and autocorrelation analysis. A successful time-series development was the maximum entropy spectral analysis of Burg [11], [12] with autoregressive (AR) models. The idea of entropy makes the autocorrelation function compatible with the given data while remaining maximally noncommittal about the data outside the observational window [11], [12]. AR modeling still required order selection, for which Akaike [13] developed the famous Akaike information criterion (AIC). The maximum-likelihood (ML) theory provides the relation between time-series models and good methods for spectral and autocorrelation estimation. The estimated AR parameters are close to the ML estimates. The invariance property of the ML theory guarantees that, asymptotically, the ML estimates of functions of parameters are just functions of the ML parameter estimates. Therefore, autocovariance and spectral estimates are considered as functions of the efficiently estimated AR parameters.

In the past, parametric AR estimation suffered from selection of overfitting models with the AIC, particularly in small samples. Moreover, the class of AR models is not generic. In addition, moving average (MA) and combined autoregressive moving average (ARMA) models are required as candidate models for arbitrary random data. No generally accepted reliable and numerically stable algorithms for MA and ARMA were available in the past [3], [4]. Moreover, the selection of the best model *type* for observed random data, i.e., AR, MA, or ARMA, had not yet been solved. Recently, reliable software has become available for the parametric time-series estimation of autocorrelation functions [14]. The spectral accuracy of the selected time-series model is mostly close to the Cramér–Rao lower bound that is attainable for the given data [10]. This gives high expectations for the quality of the autocovariance or autocorrelation function computed with a selected time-series model. The advantages of the time-series approach to autocorrelation estimation have theoretically been derived [5, p. 144]. In addition, examples have been given where the Yule–Walker method of AR estimation gives poor results [15].

Manuscript received July 4, 2007; revised November 18, 2008. First published July 28, 2009; current version published October 9, 2009. The Associate Editor coordinating the review process for this paper was Dr. Robert Gao.

The author is with the Department of Multi-Scale Physics, Delft University of Technology, 2628 CJ Delft, The Netherlands (e-mail: P.M.T.Broersen@tudelft.nl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIM.2009.2021206

This paper describes estimators for the autocovariance and autocorrelation functions, with LP or with time-series models [2], [10]. LP estimates are compared with two parametric AR approaches, namely, the Yule–Walker method [3] and the Burg method [11], [12], which is part of a general automatic time-series computer program [14]. Finite-sample relations are given between accuracy measures for autocorrelation functions and spectra and the residual variance. These can be used both for LP estimates and for AR models. LP autocorrelation functions have specific bias contributions [6], [16]. Furthermore, the overall accuracy of LP does not become better if more data are available [16]. Simulation results with stationary random processes are presented here for the accuracy of the complete autocorrelation function obtained from selected time-series models and from LP. The differences between the Yule–Walker and the Burg approach of AR estimation are demonstrated.

II. Autocorrelation Function Definitions

The definition of the autocovariance function of a stationary stochastic discrete-time process x_n, with mean zero, is for lag k given by [2]

$$r(k) = E(x_n x_{n+k}). \tag{1}$$

For wide-sense *stationary* processes, this value only depends on k and not on n. The expectations for all values of k together add up to the autocovariance function. This function is symmetric around lag $k = 0$. No value can be greater than $r(0)$, which is the variance of the signal. The normalized autocovariance function is obtained by dividing (1) by the variance of the process and is called the autocorrelation function. If the mean of the process is not known to be zero, it should be subtracted from the data before applying these definitions.

The average of the LPs $x_n x_{n+k}$ obtained from a single realization of a stochastic process might be used as an estimator for the autocorrelation function at lag k. However, the estimated function for all lags k together is not positive definite. This is a theoretical requirement for true autocorrelation functions [2]. To obtain a positive semidefinite estimator, a biased LP autocovariance has been defined, dividing the sum of $N - k$ LPs by N [2], i.e.,

$$\hat{r}_{\mathrm{LP}}(k) = \frac{1}{N} \sum_{n=1}^{N-k} x_n x_{n+k}, \quad 0 \le k < N. \tag{2}$$

The positive semidefinite LP estimator for the normalized autocorrelation function is given by

$$\hat{\rho}_{\mathrm{LP}}(k) = \frac{\sum_{n=1}^{N-k} x_n x_{n+k}}{\sum_{n=1}^{N} x_n^2} = \frac{\hat{r}_{\mathrm{LP}}(k)}{\hat{r}_{\mathrm{LP}}(0)}. \tag{3}$$

Both estimators (2) and (3) use the fact that the expectation of each product $x_n x_{n+k}$ in the summation is the true autocovariance (1). It is a historical misconception that this LP estimator has a statistical base [6]. Efficient estimators have to be derived from the statistical properties of the errors in each separate contribution, not only from the expectation in (1). As each observation x_n contributes to several LPs, the deviations in the LPs become correlated. Hence, the estimators (2) and (3) are not efficient [5]. They belong to the same category of estimators like the sum of the largest plus the smallest observation divided by 2 as an estimator for the mean value. This is another poor estimator for normal random data, which will not improve for greater data sets.

A parametric spectral or autocorrelation estimator is computed as a function of the estimated parameters of a time-series model. The time-series theory has three different model types: 1) AR; 2) MA; and 3) combined ARMA. Each stationary stochastic process can be described by a model in one of these three classes [2]. An ARMA(p, q) process x_n can be written as [2]

$$x_n + a_1 x_{n-1} + \cdots + a_p x_{n-p} = \varepsilon_n + b_1 \varepsilon_{n-1} + \cdots + b_q \varepsilon_{n-q} \tag{4}$$

where ε_n is a purely random white noise process with zero mean and variance σ_ε^2. This ARMA(p, q) process becomes AR for $q = 0$ and MA for $p = 0$. Estimates for the parameters and the model type and order can automatically be obtained with the ARMAsel program [14]. This paper deals with AR processes; therefore, only the theory of AR models is treated in the sequel.

For AR models, all lags of the infinitely long true autocorrelation function are determined by the p true AR parameters of (4). The Yule–Walker relations [3] describe the relation between the first p lags of the autocorrelation function of an AR(p) process and the parameters, i.e.,

$$\begin{aligned} &\rho_{\mathrm{TS}}(k) + a_1 \rho_{\mathrm{TS}}(k-1) + \cdots + a_p \rho_{\mathrm{TS}}(k-p) = 0, \quad k > 0 \\ &\rho_{\mathrm{TS}}(-k) = \rho_{\mathrm{TS}}(k). \end{aligned} \tag{5}$$

The Levinson–Durbin recursion is a computationally efficient recursive computation of the first p autocorrelations of an AR(p) process for given parameter values [3]. The same relation (5) can be used to compute true or estimated autocovariance functions by using true or estimated parameters, respectively.

Substitution of the biased LP estimates (3) for the autocorrelations in (5) defines the Yule–Walker method of AR estimation $(\mathrm{AR_{YW}})$ [3], [8]. Therefore, the complete estimated LP autocorrelation function from N observations is identical with the autocorrelation of the $\mathrm{AR_{YW}}(N-1)$ model. In addition, the MA$(N-1)$ model, if computed with the method of moments, would give exactly the same autocorrelation function [3]. Hence, LP autocorrelations are equal to a special representation of time-series autocorrelations with the highest possible AR or MA model orders.

The roots of the polynomial built with the AR(p) parameters as coefficients, i.e.,

$$A_p(z) = 1 + a_1 z^{-1} + \cdots + a_p z^{-p} \tag{6}$$

are denoted the poles of the AR(p) model. Always assume that data represent a stationary stochastic process, which is guaranteed if all poles of $A_p(z)$ are inside the unit circle. The power spectrum $h(\omega)$ of the AR(p) process is completely

determined by the AR parameters, together with the variance σ_ε^2, and is given by

$$h(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} \frac{1}{\left|A_p\left(e^{j\omega i}\right)\right|^2}. \tag{7}$$

The normalized spectrum $\varphi(\omega)$ is defined as the spectrum divided by the variance of the signal: $\varphi(\omega) = h(\omega)/\sigma_x^2$.

III. Accuracy Measures

An important aspect of the accuracy of autocorrelation functions is that no useful measure can be based on the autocorrelation function itself. Small differences between correlations can belong to large differences of the logarithms of spectra in weak parts of the spectral density [9]. The sum of squared differences between true and estimated spectra or autocorrelations is certainly not an adequate measure.

The spectral distortion (SD) is a relative error measure that has been defined as [10]

$$\begin{aligned} \mathrm{SD} &= \frac{0.5}{2\pi} \int_{-\pi}^{\pi} \left[\ln\{h(\omega)\} - \ln\left\{\hat{h}(\omega)\right\}\right]^2 d\omega \\ &= \frac{0.5}{2\pi} \int_{-\pi}^{\pi} \left\{\ln \frac{h(\omega)}{\hat{h}(\omega)}\right\}^2 d\omega \end{aligned} \tag{8}$$

where h is the true spectral density, and $\hat{h}$ denotes the estimated spectral density.

The cepstrum is a measure that has been defined as an accuracy measure in the time domain [8]. It is the inverse Fourier transform of the logarithm of the normalized spectral density, i.e.,

$$c(k) = \int_{-\pi}^{\pi} \ln\{\phi(\omega)\} e^{j\omega k} d\omega, \quad k = 0, \pm 1, \pm 2, \ldots. \tag{9}$$

For an AR process, the parameters a_k can be used for a recursive calculation of the cepstral coefficients $c(n)$ [8] as

$$-nc(n) - na_n = \sum_{k=1}^{n-1} (n-k)c(n-k)a_k, \quad n > 0. \tag{10}$$

The same coefficients can be found with (9). The Parseval's relation gives the equivalence in the time and the frequency domain between the cepstrum and the logarithm of the normalized spectrum, i.e.,

$$\sum_{k=-\infty}^{\infty} \{c(k) - \hat{c}(k)\}^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\ln\{\varphi(\omega)\} - \ln\{\hat{\varphi}(\omega)\}\right]^2 d\omega \tag{11}$$

where $\hat{c}(k)$ is the estimated cepstral coefficient, and $\hat{\varphi}$ is the estimated normalized spectral density. A comparison with the SD of (8) shows that the cepstral distance is almost the same as the SD, apart from a constant 2 and some variance normalization. The model with the smallest SD will also have the smallest cepstral distance.

For time series, the squared error of prediction of future data is an obvious good choice as the uncertainty measure. The squared error of prediction is often simply called the PE or the one-step-ahead PE. In simulations, where the true process is known, this can be expressed as [10]

$$\mathrm{PE} = \frac{\sigma_\varepsilon^2}{2\pi} \int_{-\pi}^{\pi} \frac{h(\omega)}{\hat{h}(\omega)} d\omega = \frac{\sigma_\varepsilon^2}{2\pi} \int_{-\pi}^{\pi} \left|\frac{\hat{A}(e^{j\omega})}{A(e^{j\omega})}\right|^2 d\omega. \tag{12}$$

The PE has a strong relation with the SD of (8) and the cepstral measure of (11). For small deviations between the true process and the estimated model parameters, they only differ by some normalization constants. For larger deviations, the global performance of the measures keeps a similarity. Accuracy measures of time series can also be used for the quality of the LP autocorrelation estimate, by using the parameters of the $\mathrm{AR_{YW}}(N-1)$ model.

Several methods exist for AR estimation [3]. The maximum entropy method of Burg [11], [12] is often used for AR estimation in practice [10]. The asymptotical theory for the residual variance s_K^2 and $\mathrm{PE}(K)$ do not depend on the estimation method. If the model order K is at least the true AR order p, the expectations for unbiased models are determined by [10]

$$E\left[s_K^2\right] = \sigma_\varepsilon^2 \left(1 - \frac{K}{N}\right), \quad K \geq p \tag{13}$$

$$E\left[\mathrm{PE}(K)\right] = \sigma_\varepsilon^2 \left(1 + \frac{K}{N}\right), \quad K \geq p. \tag{14}$$

This asymptotical result can be applied in practice with reasonable accuracy if K is less than $0.1\ N$. It might be useful to introduce a normalized PE by dividing it by σ_ε^2. However, this coincides with the PE here because σ_ε^2 will be taken as 1 in the examples.

For model orders higher than $0.1\ N$, the finite-sample theory has been introduced [10]. This theory uses empirical approximations for the variance of reflection coefficients, which are used in the Levinson–Durbin recursion [3]. The asymptotical variance $1/N$ is replaced by the variance coefficients, i.e.,

$$\begin{aligned} v_{i,\mathrm{YW}} &= \frac{N-i}{N(N+2)} \\ v_{i,\mathrm{Burg}} &= \frac{1}{N+1-i} \end{aligned} \tag{15}$$

for Yule–Walker and Burg estimates, respectively [10]. Yule–Walker variances are smaller than the asymptotical $1/N$ because of the triangular bias. In contrast, Burg variances are greater because each new Burg reflection coefficient is estimated from a shorter filtered signal [10]. The finite-sample expressions for s_K^2 and $\mathrm{PE}(K)$ have been defined as

$$E_{\mathrm{FS}}\left[s_K^2\right] = \sigma_\varepsilon^2 \prod_{i=1}^{K} (1 - v_{i,.}), \quad K \geq p \tag{16}$$

$$E_{\mathrm{FS}}\left[\mathrm{PE}(K)\right] = \sigma_\varepsilon^2 \prod_{i=1}^{K} (1 + v_{i,.}), \quad K \geq p. \tag{17}$$

By substituting the Yule–Walker or Burg variance coefficients for v_i, different finite-sample results are obtained for both estimation methods [10].

The residual variance s_K^2 is known for measured data, because it is minimized to compute the parameters. A practical approximation for the prediction accuracy of measured data can be derived with (16) and (17). It is a finite-sample equivalent of the final PE (FPE) of Akaike [13], i.e.,

$$\mathrm{FPE}_{\mathrm{FS}}(K) = s_K^2 \frac{\prod_{i=1}^{K}(1+v_{i,.})}{\prod_{i=1}^{K}(1-v_{i,.})}. \tag{18}$$

This approximation for the PE can be used for all model orders. The values of s_K^2 and $\mathrm{FPE}_{\mathrm{FS}}(K)$ are mainly determined by the true values of the reflection coefficients for the model orders $K < p$ [10]. They are independent of the true process parameters for $K \geq p$, for unbiased models.

IV. LP and AR Autocorrelations

Only the first LP autocovariances until lag $p-q$ are estimated asymptotically efficient in an ARMA(p,q) process [5]. All further LP estimates are not asymptotically efficient. This means that their variance in estimations is greater than minimally obtainable. For AR(p) processes, only the first p LP estimators are efficient and sufficient. A theoretical asymptotical analysis has been given for the accuracy of LP and AR autocorrelations of AR(1) processes [6]. The nice asymptotical theoretical properties of time-series estimates turn out to be also relevant for practical estimates from finite samples [10].

The general asymptotic expression for the covariance between the estimators of the LP autocovariance function at the lags k and $k+v$ is well known [2]. This covariance is inversely proportional to the sample size N; hence, it will finally approach zero. For larger lags k, where the true $\rho(k)$ becomes zero, an approximation for the variance of the LP estimates is given by [2]

$$\mathrm{var}\,\{\hat{\rho}_{\mathrm{LP}}(k)\} \cong \frac{1}{N}\sum_{m=-\infty}^{\infty}\rho^2(m). \tag{19}$$

This variance decreases with $1/N$, but the number of lags with nonzero LP estimates increases with the same value of N. The practical importance of this formula is that the statistical accuracy of LP estimates of decaying autocorrelation functions is a constant for all lags where the true autocorrelation $\rho(k)$ died out. Therefore, it is not possible to detect with acceptable accuracy where a correlation dies out; lifetime measurements are not well possible with LP autocorrelations. The nonzero length of the LP autocovariance function (2) is always equal to the sample size N. Taking twice as much observations roughly gives twice as much LP estimates, each with halved variance. This explains why the quality of the autocorrelation as a whole function does not improve with the sample size N but remains a constant [16].

The LP autocorrelation estimator (3) has a triangular bias and several additional bias terms [6], [16]. The LP estimator is asymptotically unbiased. Asymptotical roughly means that terms with $1/N$ disappear for increasing N and terms with $1/N^2$ can be neglected in comparison with the terms with $1/N$ for all sample sizes. The influence of the triangular bias on LP and $\mathrm{AR}_{\mathrm{YW}}$ estimates can be evaluated for a true AR(1) process with the parameter $-a$. The true autocorrelation function is $\rho_{\mathrm{TS}}(k) = a^k$, as found with (5). The biased expectation of the $\mathrm{AR}_{\mathrm{YW}}(1)$ parameter becomes $-a(1-1/N)$. The exact expectation of the biased LP estimate (3) for lag k becomes $a^k(1-k/N)$, which is asymptotically the same as $[a(1-1/N)]^k$ but has a different finite-sample value. This implies that the triangular bias transforms the true AR(1) process into an $\mathrm{AR}_{\mathrm{YW}}$ model of infinite order. The parameters may be small, but they are not identically zero. Only asymptotically, the $\mathrm{AR}_{\mathrm{YW}}$ order of a true AR(1) process will again become 1. The consequences in order selection will be investigated.

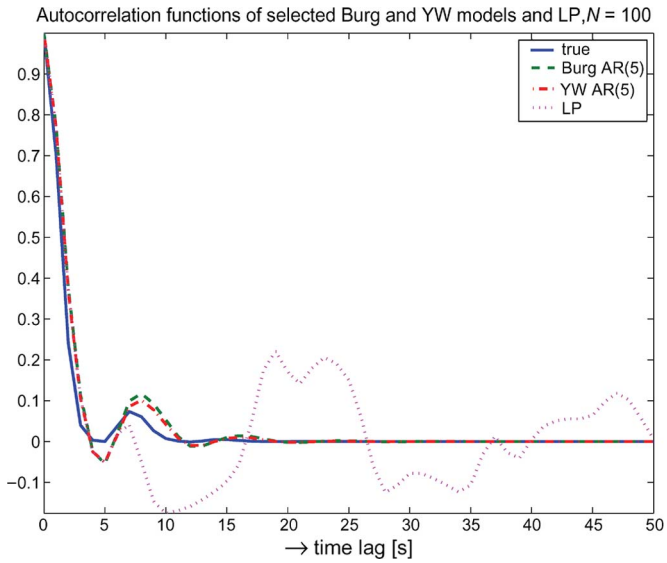


Fig. 1. True autocorrelation function of an AR(5) process with $\beta = 0.7$, the selected AR(5) Burg and YW estimates, and the LP estimate for $N = 100$. The $\mathrm{AR}_{\mathrm{YW}}(N-1)$, MA$(N-1)$, and LP estimates are identical for all lags. The PE values are 1.042, 1.034, and 1.796 for Burg, $\mathrm{AR}_{\mathrm{YW}}(5)$, and LP or $\mathrm{AR}_{\mathrm{YW}}(99)$, respectively.

Monte Carlo simulations have been done with AR(5) processes with all poles of the same radius β. All data have been generated with $\sigma_\varepsilon^2 = 1$. Therefore, normalized error measures that can be obtained by dividing by σ_ε^2 are the same as the measures in Section III. The AR(5) data have been generated by using the following true reflection coefficients:

$$(-\beta), (-\beta)^2, (-\beta)^3, (-\beta)^4, (-\beta)^5.$$

The radius and the sample size N have been varied. For N greater than about 100, the results of $\mathrm{AR}_{\mathrm{YW}}$ and AR Burg were quite close as long as β is smaller than 0.7. Fig. 1 gives the true and estimated autocorrelation functions for one realization of 100 observations, with $\beta = 0.7$. The selected orders for $\mathrm{AR}_{\mathrm{YW}}$ and AR Burg were mostly the same, in many simulation runs with this example. Their PE values were close to 1.05, which is the theoretical value of (14) for the AR(5) model. In some runs, the selected orders were 3 or 4. Sometimes, the Burg method is better, sometimes, the YW estimates give a smaller PE, but both methods give reliable results in this example.

It is possible to express the influence of the triangular bias in a PE value. The true autocorrelations at the first five lags k are multiplied with $1 - k/N$, and the AR parameters are computed with these biased correlations substituted in the Yule–Walker equations (5). The PE value of the theoretical triangular bias with $\beta = 0.7$ is 1.0090 for $N = 100$. It would become 1.00012 for $N = 1000$. Therefore, the bias contribution is much smaller than the estimation uncertainty of (14). For these situations, the Burg and Yule–Walker estimates are similar if order selection is used. The accuracy of the complete LP autocorrelation function of length $N - 1$ is always poor for all examples.

It is not possible to extract or approximate the PE values with the information available in Fig. 1. Only large differences between true and estimated autocorrelations can, with certainty, be attributed to poor estimates; small differences might belong to either good or poor estimates. This influence of small differences can be demonstrated as follows: The LP estimate has PE $= 1.796$ here. The estimated LP autocorrelation for lag 99 was -0.0040. The periodogram is the inverse Fourier transform of the LP and is positive for all frequencies, as it should be. The inverse Fourier transform has also been computed for the LP autocorrelation with only the estimate for the highest lag 99 equated to zero. The inverse transform was now negative at 5 of the 100 FFT frequencies. The SD of (8) and the PE are computed with the logarithm of the spectrum, which does not anymore exist for negative spectral estimates. The very small variation of the LP autocorrelation function changes the PE from a finite to an infinite number. The same effect of negative spectral estimates can be obtained by adding or subtracting a small amount from LP estimates at arbitrary lags.

The LP estimate does not die out in Fig. 1 because of (19). The true and estimated AR autocorrelations are extinct for lags greater than 15 or 20. LP and $\mathrm{AR_{YW}}(5)$ are identical for the first five lags and become different for greater lags. It is clear that the LP estimates cannot be used to estimate the length of the autocorrelation function in this example. It has already been demonstrated with 1000 observations of an AR(4) example that LP estimates are not extinct [16]. With (19), the squares of all nonzero autocorrelations contribute to the LP variance at greater lags. This prohibits the search for lags where the correlation stays at values less than a value such as $1/\sqrt{N}$. Hence, the nonzero length of LP estimates is $N - 1$. It is determined by the length of the observation window and not by the character of the data [16].

The estimated model accuracy can be computed for all model orders, for Yule–Walker and Burg estimates. They are given in Fig. 2, together with their finite-sample expectations. It is clear that both methods have different results in finite samples. The finite-sample theory always gives close results for all AR methods if orders are lower than $N/10$, for all N. The estimated accuracy of model orders lower than 5 is poor for an AR(5) example, because (17) does not include true process parameters. The finite-sample expectation has been computed with (17) for all model orders, as if the true process were white noise. The asymptotical theory of (14) would give a straight line to the PE value 2 for order 100, for both methods. This agrees with the lines in Fig. 2 for low orders and is between the given finite-sample lines for higher orders. It has been shown in

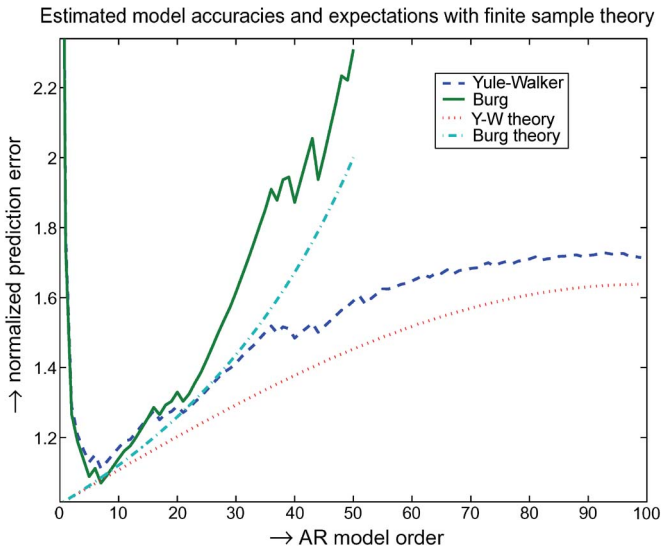


Fig. 2. Estimated $\mathrm{FPE_{FS}}$ of (18) and the expectation $E_{\mathrm{FS}}[\mathrm{PE}]$ of (17) as a function of the model order for Yule–Walker and Burg estimates. Results of the single representative simulation run of Fig. 1 are given until order $N - 1$ for Yule–Walker estimates and until $N/2$ for Burg. The accuracy of Yule–Walker for order $N - 1$ is the LP accuracy.

many examples that the average of many runs converges to the theoretical expectations for model orders higher than the true order [10]. The expectation of the accuracy PE of the LP estimate for order $N - 1$ is computed for unbiased models. This is rather accurate here, because the triangular bias is negligible. The PE quality of LP estimates is about 1.65, independent of N. The triangular bias of the Yule–Walker estimates is the cause of the inaccuracy of higher orders being somewhat better than the Burg inaccuracy. However, the accuracy of the selected models of order 5 is about the same.

The accuracy of LP is the worst of all AR candidates of the Yule–Walker method, except that of orders lower than 5. Therefore, order selection between the Yule–Walker candidates gives an important improvement of the quality of the model. With the theoretical invariance property of the ML theory, this applies to its autocorrelation function and estimated spectral density. It has been shown in numerous simulations that this property may be applied to practical time-series estimates of selected models obtained with the ARMASA toolbox [14]. This toolbox uses the Burg method for AR estimation [10].

V. Influential Triangular Bias

Asymptotically, the influence of the triangular bias will always disappear. However, for $N = 100$, it becomes important in an AR(5) process with five poles at equal radii if the radius β is greater than 0.8. Fig. 3 gives the true poles of the AR(5) process with $\beta = 0.85$ and the poles of the true biased process, computed with the expectation of the triangular bias for the first five lags of the autocorrelation function. The biased poles are far away from the true poles, particularly the pole on the real axis. The PE of the true biased model, without estimation uncertainty, is 1.619 for $N = 100$, and it would become 1.039 for $N = 1000$.

Table I gives the true and biased reflection coefficients that belong to the poles in Fig. 3. The first κ_b has a bias of $1/N$, the

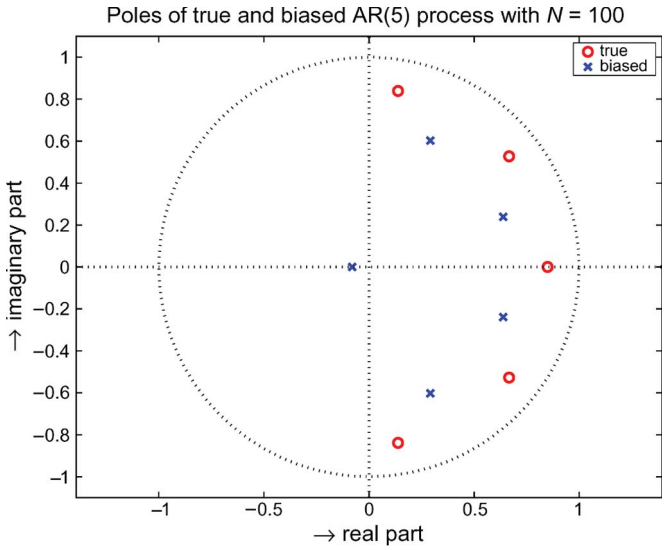


Fig. 3. True poles of the AR(5) process with $\beta = 0.85$ and the poles after applying the triangular bias to the first five lags of the true correlation function, for $N = 100$. The PE of the biased true process is 1.619.

TABLE I
TRUE REFLECTION COEFFICIENTS κ AND BIASED REFLECTION COEFFICIENTS κ_b THAT BELONG TO THE POLES OF FIG. 3

order	κ	κ_b
1	-0.85	-0.84
2	0.72	0.67
3	-0.61	-0.44
4	0.52	0.17
5	-0.44	0.02

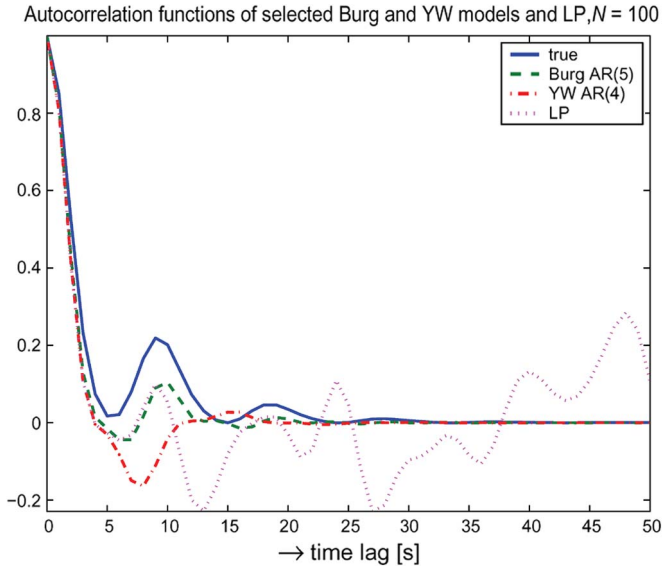


Fig. 4. True autocorrelation function of an AR(5) process with $\beta = 0.85$, the selected AR(5) Burg and AR(4) YW estimates, and the LP estimate for $N = 100$. The PE values in this realization are 1.026, 1.359, and 2.173 for Burg, $\mathrm{AR}_{\mathrm{YW}}(4)$, and LP or $\mathrm{AR}_{\mathrm{YW}}(99)$, respectively.

fourth κ_b is only 1/3 of the true value, and the fifth κ_b has the wrong sign. For increasing sample sizes, the biased poles will converge to true values.

Fig. 4 gives the true and some estimated autocorrelation functions. The LP estimate remains irregular at greater lags, as expected. The AR estimates with the Burg and Yule–Walker methods converge to zero. Although it is not possible to derive an objective accuracy measure from this figure, it is obvious that the character of LP is completely different from the true process. In repeated simulations, the Burg methods selected order 5 in the majority of runs. A lower order has never been selected, and only in about 10% of the runs have order 6 been selected. The average PE of the selected Burg model was about 1.05. The selected order and the PE of the Yule–Walker method varied much more: The order was between 2 and 7, and the PE was between 1.2 and 2.

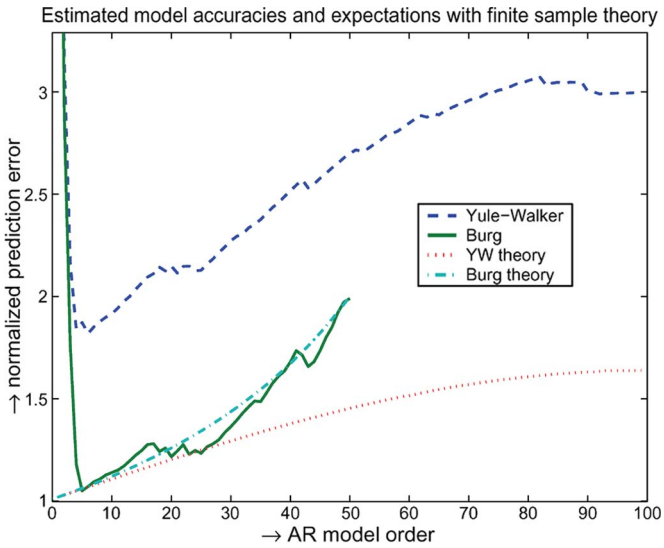


Fig. 5. Estimated $\mathrm{FPE}_{\mathrm{FS}}$ of (18) and its expectation $E_{\mathrm{FS}}[\mathrm{PE}]$ of (17), both as a function of the model order. The results of the single run of Fig. 4 are given. The bias of $\mathrm{AR}_{\mathrm{YW}}$ has an effect on all model orders and is not diminished or compensated by using higher model orders.

Fig. 5 shows the estimated model accuracy and the expectations that have been derived for unbiased estimates. The accuracy of AR Burg estimates is comparable with that in Fig. 2. The inaccuracy of $\mathrm{AR}_{\mathrm{YW}}$ is much higher than the unbiased expectation. The numerical values of the different error measures are no longer exactly the same for the large errors involved. The triangular bias of the Yule–Walker estimates has two visible consequences in Fig. 2: First, the level of the inaccuracy is much higher than the theoretical unbiased level for all model orders. Second, the slope at high frequencies flattens out because of the large triangular bias at lags close to N. This is also visible in the finite-sample theory. In contrast, the slope of Burg estimates becomes steeper at higher orders.

The true and estimated power spectral densities are given in Fig. 6 for the 100 observations that have been used. The selected AR Burg spectrum is close to the true spectrum. The selected $\mathrm{AR}_{\mathrm{YW}}$ spectrum roughly follows the biased expectation of the spectrum. The periodogram as inverse transform of the LP has the same bias and would also be around the biased expectation, with much larger variations.

If a linear plot of the spectrum would be presented, all four lines would almost be completely shrunken to the axis for frequencies above 0.25 Hz, and they would have the same value of zero there. The differences seem to be small in linear plots. To emphasize the importance of the differences in the high-frequency part of the spectrum, adequate measures for spectra

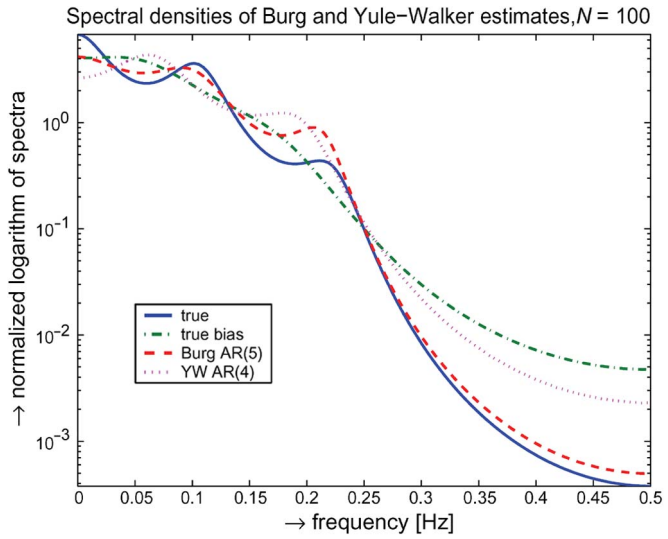


Fig. 6. True power spectral density of the realization of an AR(5) process with $\beta = 0.85$, the biased expectation, and the selected AR(5) Burg and AR(4) YW estimates for $N = 100$. The simulation run of Fig. 4 has been used.

are based on the logarithms or the quotients of spectra. It has been verified that the four spectra in Fig. 6 coincide within the line width if $N = 10\,000$ observations of the same true process are used. In addition, the true and estimated accuracy in Fig. 5 would be very close then.

The autocorrelations selected with the AR method of Burg can only become inaccurate if not enough data are available to obtain accurate estimates. Furthermore, an occasional small estimate of the last parameter may sometimes give a lower order with order selection. This can happen with statistical criteria. LP and $\mathrm{AR_{YW}}$ autocorrelations, however, cannot be accurate either in these circumstances. LP and $\mathrm{AR_{YW}}$ estimates can become inaccurate for very significant details, due to the bias. Hence, examples can be given where LP and $\mathrm{AR_{YW}}$ autocorrelation estimates do not represent the characteristics of the data, whereas the AR Burg autocorrelation function shows all relevant details.

VI. Concluding Remarks

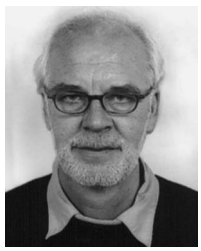
Although the triangular bias in the autocorrelation function is of the magnitude $1/N$, the influence of that small bias on the estimated characteristics of the data can be very significant. LP, periodograms, and AR estimates of the Yule–Walker method suffer from this bias, whereas the AR method of Burg is free of it. Therefore, the Burg's method is a better choice as part of an automatic analysis of measured random data with unknown characteristics.

The LP estimate for the autocorrelation function is the same as the autocorrelation of the $\mathrm{AR_{YW}}(N-1)$ model. The accuracy of the LP estimates for the autocorrelation function as a whole does not become better if more data are available. Only the accuracy for single lags improves. Selection of the best order between all $\mathrm{AR_{YW}}$ candidates gives a better estimate of the autocorrelation function than LP. It can be of the same quality as a selected AR model estimated with the Burg's method, but only if triangular bias is negligible. The influence of the triangular bias can be evaluated by comparing the PE of the biased AR model with the expected PE of estimated true order models.

The lifetime of a process can be defined as the length of the autocorrelation function. This can be estimated from time-series correlations and not from LPs.

References

[1] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Math. Comput.*, vol. 19, no. 90, pp. 297–301, Apr. 1965.
[2] M. B. Priestley, *Spectral Analysis and Time Series*. London, U.K.: Academic, 1981.
[3] S. M. Kay and S. L. Marple, "Spectrum analysis—A modern perspective," *Proc. IEEE*, vol. 69, no. 11, pp. 1380–1419, Nov. 1981.
[4] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Upper Saddle River, NJ: Prentice–Hall, 1997.
[5] B. Porat, *Digital Processing of Random Signals*. Englewood Cliffs, NJ: Prentice–Hall, 1994.
[6] P. M. T. Broersen, "Historical misconceptions in autocorrelation estimation," *IEEE Trans. Instrum. Meas.*, vol. 56, no. 4, pp. 1189–1197, Aug. 2007.
[7] B. Hof, J. Westerweel, T. M. Schneider, and B. Eckhardt, "Finite lifetime of turbulence in shear flows," *Nature*, vol. 443, no. 7107, pp. 59–62, Sep. 7, 2006.
[8] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 63, no. 4, pp. 561–580, Apr. 1975.
[9] P. M. T. Broersen, "The performance of spectral quality measures," *IEEE Trans. Instrum. Meas.*, vol. 50, no. 3, pp. 813–818, Jun. 2001.
[10] P. M. T. Broersen, *Automatic Autocorrelation and Spectral Analysis*. London, U.K.: Springer-Verlag, 2006.
[11] J. P. Burg, "Maximum entropy spectral analysis," in *Proc. 37th Meeting Soc. Explor. Geophys.*, 1967, pp. 1–6.
[12] *Modern Spectrum Analysis*, IEEE Press, New York, 1978, pp. 34–41.
[13] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. AC-19, no. 12, pp. 716–723, Dec. 1974.
[14] P. M. T. Broersen, *The ARMASA MATLAB Toolbox*, 2002. (signal processing, spectral analysis). [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange
[15] M. J. L. de Hoon, T. H. J. J. van der Hagen, H. Schoonewelle, and H. van Dam, "Why Yule–Walker should not be used for autoregressive modelling," *Ann. Nucl. Energy*, vol. 23, no. 15, pp. 1219–1228, 1996.
[16] P. M. T. Broersen, "Length and quality of lagged-product autocorrelation estimates," in *Proc. IEEE/IMTC Conf.*, Warsaw, Poland, May 2007, pp. 1–6.



Piet M. T. Broersen was born in Zijdewind, The Netherlands, in 1944. He received the M.Sc. degree in applied physics and the Ph.D. degree from Delft University of Technology, Delft, The Netherlands, in 1968 and 1976, respectively.

He is currently with the Department of Multi-Scale Physics, Delft University of Technology. He developed statistical measures to let the measured data speak for themselves in estimated time-series models. This provides a practical and accurate solution for the spectral and autocorrelation analysis of stochastic data. Software for the automatic estimation of spectra and autocorrelation functions of equidistant random data, for missing data problems, and for irregularly sampled observations, is available online. His main research interest is in the automatic and unambiguous identification of the character of stationary random measurement data.