UNBIASED EQUATION ERROR IDENTIFICATION AND APPROXIMATIONS: A FREQUENCY-DOMAIN SOLUTION AND ITS PERFORMANCE ANALYSIS

Channarong Tontiruttananon & Jitendra K. Tugnait
Department of Electrical Engineering
Auburn University, Auburn, Alabama 36849, USA
tugnait@eng.auburn.edu

ABSTRACT

We consider a frequency-domain solution to the least-squares equation error identification problem using the power spectrum and the cross-spectrum of the IO (input-output) data to estimate the IO parametric transfer function. The proposed approach is shown to yield a unimodal performance surface, consistent identification in colored noise and sufficient-order case, and stable fitted models under undermodeling for arbitrary stationary inputs so long as they are persistently exciting of sufficiently high order. Asymptotic performance analysis is carried out for both sufficient-order and reduced-order cases. Computer simulation results are presented to illustrate the proposed approach.

1 Introduction

Consider the following widely used input-output linear system model:

$$y(t) = H(q)u(t) + v(t) \qquad (1-1)$$

where $\{u(t)\}$ is the measured input sequence, t is discrete-time, $\{y(t)\}$ is the noisy output, and $\{v(t)\}$ is a measurement noise (disturbance) sequence. With q^{-1} denoting the backward-shift operator (i.e. $q^{-1}u(t) = u(t-1)$), the linear system $H(q)$ represents an IIR (infinite impulse response) system:

$$H(q) = \sum_{i=0}^{\infty} h(i)q^{-i}. \qquad (1-2)$$

Given an input-output record $\{u(t), y(t), \quad t = 1, 2, \cdots\}$, but the underlying true system model $H(q)$ ***unknown***, it is of much interest in control, communications and signal processing applications to fit a rational function model

$$G(q) := \frac{B(q)}{A(q)} = \frac{\sum_{i=1}^{n_b} b_i q^{-i}}{1 + \sum_{i=1}^{n_a} a_i q^{-i}} \qquad (1-3)$$

to given input-output record [1]-[6],[8]. A wide variety of approaches exist [1],[4],[5],[8].

The main objective of this paper is to provide a frequency-domain solution using spectral analysis to the problem of equation error (least-squares) system identification given time-domain input-output data. The proposed method is shown to lead to:

This work was supported by the National Science Foundation under Grant ECS-9504878.

- Global convergence (unimodal cost function) since we have a cost quadratic in the unknown parameters. This is unlike PEM (prediction error method) and OEM (output error method) [4],[5].
- Consistent estimates in the sufficient order case even when $\{v(t)\}$ is colored. This is unlike the EEM (equation error method) of [1] which requires the noise to be white.
- Asymptotically stable fitted model in the reduced order case for arbitrary stationary inputs so long as they are persistently exciting of sufficiently high order. This is unlike PEM, SSM (Steiglitz-McBride method), EEM and IVM (instrumental variable method). In particular, ARMA inputs are included unlike [1].

2 Model Assumptions

We impose the following conditions on (1-1):

(AS1) $\{u(t)\}$ and $\{y(t)\}$ are zero-mean and jointly stationary. The power spectral density (PSD) $S_{uu}(\omega)$ of $\{u(t)\}$ is > 0 for almost all $\omega \in [0, \pi]$ if the proposed approaches utilize the entire frequency range $[0, \pi]$. If a finite number of frequencies are used then $S_{uu}(\omega)$ need be nonzero only for this frequency set.

(AS2) The true system transfer function $H(q)$ is causal and stable. Therefore, $\sum_{i=0}^{\infty} h^2(i) < \infty$.

(AS3) The noise sequence $\{v(t)\}$ is zero-mean, stationary and independent of $\{u(t)\}$.

(AS4) The following summability conditions hold true:

$$\sum_{\tau_1, \cdots, \tau_{k-1} = -\infty}^{\infty} [1+|\tau_j|] |C_{z_1 z_2 \cdots z_k}(\tau_1, \cdots, \tau_{k-1})| < \infty,$$

for each $j = 2, \cdots, k-1$ and each $k = 2, 3, \cdots$ where $z_i(t) \in \{y(t), u(t), v(t)\}$ and $C_{z_1 z_2 \cdots z_k}(\tau_1, \cdots, \tau_{k-1})$ is the k-th order joint cumulant of the random variables $\{z_1(t + \tau_1), \cdots, z_{k-1}(t + \tau_{k-1}), z_k(t)\}$.

Let the vector of unknown parameter be given by

$$\theta = [a_1 \ \cdots \ a_{n_a} \ b_0 \ \cdots \ b_{n_b}]^T. \qquad (2-1)$$

3 A Frequency-Domain Solution

Consider the cross-spectral density

$$S_{yu}(\omega) = \sum_{k=-\infty}^{\infty} E\{y(t+k)u(t)\}e^{-j\omega k}. \qquad (3-1)$$

It then follows easily that

$$H(e^{j\omega}) = H(q)|_{q=e^{j\omega}} = \frac{S_{yu}(\omega)}{S_{uu}(\omega)}. \qquad (3-2)$$

The basic approach to model parameter estimation consists of two steps. First obtain a consistent estimator $\widehat{H}(e^{j\omega})$ of $H(e^{j\omega})$ via consistent estimators $\widehat{S}_{yu}(\omega)$ and $\widehat{S}_{uu}(\omega)$ of $S_{yu}(\omega)$ and $S_{uu}(\omega)$, respectively, based upon the input-output record $\{u(t), y(t),\ t = 1, 2, \cdots, T\}$. Next estimate the system parameters using the estimated transfer function matrix as "data."

3.1 Transfer Function Estimator

Given a record of length T, let $Y(\omega)$ denote the DFT of $\{y(t),\ 1 \le t \le T\}$ given by

$$Y(\omega_k) = \sum_{t=0}^{T-1} y(t+1)\exp(-j\omega_k t), \qquad (3-3)$$

$$\omega_k = \frac{2\pi}{T}k, \quad k = 0, 1, \cdots, T-1. \qquad (3-4)$$

Similarly define $U(\omega_k)$. Given the above DFT's, following [7, Sec. 7.4] we define the cross- and auto-spectrum estimators as

$$\widehat{S}_{yu}(k) = \frac{1}{T(2m_T+1)} \sum_{i=-m_T}^{m_T} Y(\omega_{k-i})U^*(\omega_{k-i}), \quad (3-5)$$

$$\widehat{S}_{uu}(k) = \frac{1}{T(2m_T+1)} \sum_{i=-m_T}^{m_T} U(\omega_{k-i})U^*(\omega_{k-i}). \quad (3-6)$$

Let us choose m_T to be such that as $T \to \infty$, we have $m_T T^{-1} \to 0$ and $m_T \to \infty$. Let $k_l(T)$ with $T = 1, 2, \cdots$ be a sequence of integers such that $\lim_{T\to\infty} k_l(T)/T = f_l$, a fixed frequency (in Hz). In light of (3-5) define a coarser frequency grid:

$$\omega_l = \frac{2\pi l}{L_T} = \frac{2\pi l(2m_T+1)}{T} = \frac{2\pi l(2B_T T+1)}{T} \quad (3-7)$$

with $l = 0, 1, \cdots, L_T - 1$ where $L_T = \lfloor \frac{T}{2m_T+1} \rfloor$. Using the estimated spectra we have an estimator of the system transfer function at frequency ω_k (as in [7, Chapter 8])

$$\widehat{H}(e^{j\omega_k}) = \widehat{S}_{uu}^{-1}(k)\widehat{S}_{yu}(k) \qquad (3-8)$$

provided that $\widehat{S}_{uu}^{-1}(k)$ exists. If $S_{uu}^{-1}(\omega_k)$ exists, then it follows from [7, Thm. 8.11.1] that

$$\lim_{T\to\infty} \widehat{H}(e^{j2\pi f}) = \lim_{T\to\infty} \widehat{S}_{uu}^{-1}(k(T))\widehat{S}_{yu}(k(T)) = H(e^{j2\pi f}) \ \text{ w.p.1} \qquad (3-9)$$

where $\lim_{T\to\infty} k(T)/T = f$. Convergence in (3-9) is uniform in f.

As before, let $k_l(T)$ with $T = 1, 2, \cdots$ be a sequence of integers such that $\lim_{T\to\infty} k_l(T)/T = f_l$. We may take these integers to belong to the coarser grid $\{k \mid k = l(2m_T+1),\ l = 0, 1, \cdots, (L_T/2)-1\}$. Consider a fixed set of M frequencies λ_l for $l = 1, 2, \cdots, M$ such that $0 \le \lambda_1 < \lambda_2 < \cdots < \lambda_M < \pi$, where $\lambda_l = 2\pi f_l$. It follows from [7, Thm. 8.8.1] (see also [7, Thm. 7.4.3] and [7, Cor. 7.4.3]) that, for large T, $\widehat{H}(e^{j\lambda_l})$ for $l = 1, 2, \cdots, M$ are (asymptotically) jointly complex (circularly symmetric) Gaussian such that for large T

$$\mathrm{cov}\left(\widehat{H}(e^{j\lambda_k}), \widehat{H}(e^{j\lambda_l})\right) = \Delta_T^{-1}\sigma^2(\lambda_k)\delta(k-l) + O(T^{-1}), \qquad (3-10)$$

$$\mathrm{cov}\left(\widehat{H}(e^{j\lambda_k}), \widehat{H}^*(e^{j\lambda_l})\right) = O(T^{-1}) \qquad (3-11)$$

where $\Delta_T = 2m_T + 1$,

$$\sigma^2(\lambda_k) := \frac{S_{yy}(\lambda_k)}{S_{uu}(\lambda_k)}\left[1 - \frac{|S_{yu}(\lambda_k)|^2}{S_{yy}(\lambda_k)S_{uu}(\lambda_k)}\right]. \quad (3-12)$$

and $\mathrm{cov}\{X, Y\} = E\{XY^*\} - E\{X\}E\{Y^*\}$. Thus, $\widehat{H}(e^{j\omega_k})$ on the coarse grid (3-7) is asymptotically a complex Gaussian (in the sense of [7, Sec. 4.2]) random variable, independent at distinct frequencies on the coarse grid over $(0, \pi)$, with the covariance structure (3-10).

Remark 1. In the rest of the paper we will use ω_k to denote a frequency on the coarse grid (3-7) with $k = 0, 1, \cdots, L_T - 1$ but we will use λ_k to denote a fixed frequency independent of the record length T. •

3.2 An Equation Error Formulation

Choose θ to minimize the cost

$$J_{1T}(\theta; M, \Omega_L, \Omega_U) = \frac{1}{M}\sum_{l=1}^{M} \left|A(e^{j\lambda_l};\theta)\widehat{H}(e^{j\lambda_l}) - B(e^{j\lambda_l};\theta)\right|^2 \qquad (3-13)$$

where $0 \le \Omega_L \le \lambda_1 < \lambda_2 < \cdots < \lambda_M \le \Omega_U \le \pi$,

$$B(e^{j\lambda_l};\theta) = \sum_{i=1}^{n_b} b_i(\theta)e^{-j\lambda_l i}, \qquad (3-14)$$

$$A(e^{j\lambda_l};\theta) = 1 + \sum_{i=1}^{n_a} a_i(\theta)e^{-j\lambda_l i}. \qquad (3-15)$$

Proof of the following result is omitted.

Lemma 1. (A) Under (AS1)-(AS4) such that $0 \le \Omega_L \le \lambda_1 < \lambda_2 < \cdots < \lambda_M \le \Omega_U \le \pi$, $\lim_{T\to\infty} J_{1T}(\theta; M, \Omega_L, \Omega_U) \overset{w.p.1}{=} J_{1\infty}(\theta; M, \Omega_L, \Omega_U)$ uniformly in θ for $\theta \in \Theta_C$, any compact set, where

$$J_{1\infty}(\theta; M, \Omega_L, \Omega_U) = \frac{1}{M}\sum_{l=1}^{M} \left|A(e^{j\lambda_l};\theta)H(e^{j\lambda_l}) - B(e^{j\lambda_l};\theta)\right|^2 \qquad (3-16)$$

(B) In addition, let the set of frequencies above become dense in the interval $[\Omega_L, \Omega_U]$ as $M \to$

∞ where λ_i's are spaced uniformly in this interval. Then $\lim_{M\to\infty}\lim_{T\to\infty} J_{1T}(\theta; M, \Omega_L, \Omega_U) \overset{w.p.1}{=} J_{1\infty}(\theta; \infty, \Omega_L, \Omega_U)$ uniformly in θ for $\theta \in \Theta_C$ where

$$J_{1\infty}(\theta; \infty, \Omega_L, \Omega_U)$$

$$= \int_{\Omega_L}^{\Omega_U} \left| A(e^{j\omega}; \theta) H(e^{j\omega}) - B(e^{j\omega}; \theta) \right|^2 \frac{d\omega}{\pi} \qquad (3-17)$$

Remark 2. Suppose that we had access to noise-free measurements

$$x(t) := y(t) - v(t) = H(q)u(t). \qquad (3-18)$$

Consider the time-domain least-squares parameter estimation problem where we fit model $G(q)$ to data $\{u(t), x(t)\}$ [4, Sec. 7.1]. Choose θ to minimize $E\{\epsilon^2(t)\}$ where

$$\epsilon(t) := x(t) + \sum_{i=1}^{n_a} a_i x(t-i) - \sum_{i=1}^{n_b} b_i u(t-i). \qquad (3-19)$$

It has been established in [1] that

$$E\{\epsilon^2(t)\} = \int_{-\pi}^{\pi} \left| A(e^{j\omega}; \theta) H(e^{j\omega}) - B(e^{j\omega}; \theta) \right|^2 S_{uu}(\omega) \frac{d\omega}{2\pi} \qquad (3-20)$$

If $\{u(t)\}$ is white with variance σ_u^2 then the right-side of (3-20) equals $J_{1\infty}(\theta; \infty, 0, \pi)$ to within a scale factor. That is, asymptotically as both T and $M \to \infty$ (see also Theorem 1 in Sec. 4), *minimization of* (3-13) *yields the same mean* estimator and model fit as would have been obtained if the true system (1-1) were driven by *white* sequence $\{u(t)\}$ and noise-free measurements were available. This then is the main justification for seeking a frequency-domain solution given time-domain data. It is known that under noise-free measurements and white input, the least-squares solution has some very attractive properties [9]; these are discussed in Sec. 4. □

4 Convergence Analysis

Define

$$\widehat{\theta}_{TM}^{(1)} = \arg\{\min_\theta J_{1T}(\theta; M, \Omega_L, \Omega_U)\}, \qquad (4-1)$$

$$\overline{\theta}_{M}^{(1)} = \arg\{\min_\theta J_{1\infty}(\theta; M, \Omega_L, \Omega_U)\}, \qquad (4-2)$$

$$\overline{\theta}^{(1)} = \arg\{\min_\theta J_{1\infty}(\theta; \infty, \Omega_L, \Omega_U)\}. \qquad (4-3)$$

Theorem 1. Under the hypotheses of Lemma 1, it follows that

$$\lim_{T\to\infty} \widehat{\theta}_{TM}^{(1)} \overset{w.p.1}{\in} \mathcal{D}_M^{(1)}(\Omega_L, \Omega_U)$$

$$:= \left\{ \theta \;\middle|\; J_{1\infty}(\theta; M, \Omega_L, \Omega_U) = J_{1\infty}(\overline{\theta}_M^{(1)}; M, \Omega_L, \Omega_U) \right\}, \qquad (4-4)$$

$$\lim_{M\to\infty}\lim_{T\to\infty} \widehat{\theta}_{TM}^{(1)} \overset{w.p.1}{\in} \mathcal{D}_\infty^{(1)}(\Omega_L, \Omega_U)$$

$$:= \left\{ \theta \;\middle|\; J_{1\infty}(\theta; \infty, \Omega_L, \Omega_U) = J_{1\infty}(\overline{\theta}^{(1)}; \infty, \Omega_L, \Omega_U) \right\} \bullet \qquad (4-5)$$

Proof: Mimic the proof of Theorem 1 in [10] using Lemma 1. Note that the convergence to the set $\mathcal{D}^{(1)}$ is to be interpreted in the sense of Ljung [5, p. 215]. □

The properties of $\overline{\theta}^{(1)}$ for $\Omega_L = 0$ and $\Omega_U = \pi$ have been studied in [9]. First we need some definitions.

Def. **Sufficient Order Model Set:** The true model $H(q)$ is of the type (1-3) such that the true model orders n_{a0} and n_{b0} satisfy $\min(n_a - n_{a0}, n_b - n_{b0}) \geq 0$. •

Def. **Reduced Order Model Set (Undermodeling):** Either the true model $H(q)$ is not of the type (1-3), or it is but the true model orders n_{a0} and n_{b0} satisfy $\min(n_a - n_{a0}, n_b - n_{b0}) < 0$. •

It has been shown in [9] that under the **sufficient order** case, $\mathcal{D}_\infty^{(1)}(0, \pi)$ equals the set

$$\mathcal{D}^{(so)} := \{\theta \mid B(q;\theta)/A(q;\theta) = H(q)\}. \qquad (4-6)$$

Under **undermodeling** and $\Omega_L = 0$ and $\Omega_U = \pi$, by [9, Prop. 2], the zeros of $A(q; \overline{\theta}^{(1)})$ lie in the open unit disk; hence the fitted model $\widehat{G}(q) = B(q; \overline{\theta}^{(1)})/A(q; \overline{\theta}^{(1)})$ is stable. Moreover, under undermodeling, $\Omega_L = 0$ and $\Omega_U = \pi$, $\overline{\theta}^{(1)}$ is unique (i.e. $\mathcal{D}_\infty^{(1)}(0, \pi)$ is a singleton), and $J_{1\infty}(\overline{\theta}^{(1)}; 0, \pi) > 0$.

Using the above results from [9] and Theorem 1, the following result is immediate.

Theorem 2. Under the hypotheses of Lemma 1, $\Omega_L = 0$, $\Omega_U = \pi$ and undermodeling,

$$\lim_{M\to\infty}\lim_{T\to\infty} \widehat{\theta}_{TM}^{(1)} \overset{w.p.1}{=} \overline{\theta}^{(1)}$$

where $\overline{\theta}^{(1)}$ is unique such that the zeros of $A(q; \overline{\theta}^{(1)})$ lie in the open unit disk; hence the fitted model $\widehat{G}(q) = B(q; \overline{\theta}^{(1)})/A(q; \overline{\theta}^{(1)})$ is stable. Moreover, $\widehat{g}(i) = h(i)$ for $i = 0, 1, \cdots, n_b$ where $\widehat{G}(q) = \sum_{i=0}^{\infty} \widehat{g}(i) q^{-i}$ and $h(i)$ is as in (1-2). Under (AS1)-(AS4), $\Omega_L = 0$, $\Omega_U = \pi$ and sufficient order modeling,

$$\lim_{M\to\infty}\lim_{T\to\infty} \widehat{\theta}_{TM}^{(1)} \overset{w.p.1}{\in} \mathcal{D}^{(so)}.$$

If $\min(n_a - n_{a0}, n_b - n_{b0}) = 0$, then $\mathcal{D}^{(so)}$ is a singleton. •

Using [3, Lemma 5] and Theorem 1, the following result is immediate.

Theorem 3. Under (AS1)-(AS4), $\Omega_L > 0$, $\Omega_U < \pi$ and sufficient order modeling such that $n_a + n_b \leq 2M$, it follows that

$$\lim_{T\to\infty} \widehat{\theta}_{TM}^{(1)} \overset{w.p.1}{\in} \mathcal{D}^{(so)}.$$

If $\min(n_a - n_{a0}, n_b - n_{b0}) = 0$, then $\mathcal{D}^{(so)}$ is a singleton. •

5 Performance Analysis

We now state some results without any proofs. We will use the short notation $J_{1T}'(\theta)$ for $J_{1T}(\theta; M, \Omega_L, \Omega_U)$. Also we will use $\overline{\theta}^{(1)}$ for both $\overline{\theta}_M^{(1)}$ and $\overline{\theta}^{(1)}$, it being clear from context as to which is meant. Let ∇_θ denote the gradient operator w.r.t. vector θ. Similarly denote the Hessian matrix by $\nabla_{\theta\theta}^2 J$ whose ij-th element is $\frac{\partial^2 J}{\partial\theta_i \partial\theta_j}$.

It can be shown that $\widehat{\theta}^{(1)}_{TM}$ is asymptotically Gaussian with mean $\overline{\theta}^{(1)}$ and

$$\mathrm{cov}\left\{\widehat{\theta}^{(1)}_{TM}, \widehat{\theta}^{(1)}_{TM}\right\}$$

$$= \frac{L_T}{TM}\left[\nabla^2_{\theta\theta} J'_{1T}(\overline{\theta}^{(1)})\right]^{-1} \Sigma_{\theta M}\left[\nabla^2_{\theta\theta} J'_{1T}(\overline{\theta}^{(1)})\right]^{-1} + O(T^{-1}) \qquad (5-1)$$

where

$$\nabla^2_{\theta\theta} J'_{1T}(\overline{\theta}^{(1)}) \overset{T\to\infty}{=} \frac{1}{M}\sum_{l=1}^{M}\left(\overline{C}_l\overline{C}_l^{\mathcal{H}} + \overline{C}_l^{*}\overline{C}_l^{T}\right) \text{ w.p.1,}$$

$$\overline{C}_l = \left[e^{j\lambda_l} H^*(e^{j\lambda_l}) \vdots \cdots \vdots e^{jn_a\lambda_l} H^*(e^{j\lambda_l}) \vdots -e^{j\lambda_l} \vdots \cdots \vdots -e^{jn_b\lambda_l}\right]^T$$

$$\Sigma_{\theta M} = \frac{1}{M}\sum_{l=1}^{M}\sigma^2(\lambda_l)\left(\mathcal{F}_l\mathcal{F}_l^{\mathcal{H}} + \mathcal{F}_l^{*}\mathcal{F}_l^{T}\right),$$

$$\mathcal{F}_l = A(e^{j\lambda_l};\overline{\theta}^{(1)})\overline{C}_l + \overline{\mathcal{E}}_l(\overline{\theta}^{(1)})\mathcal{C}^*_{lg},$$

$$\overline{\mathcal{E}}_l(\overline{\theta}^{(1)}) = A(e^{j\lambda_l};\overline{\theta}^{(1)})H(e^{j\lambda_l}) - B(e^{j\lambda_l};\overline{\theta}^{(1)}),$$

$$\mathcal{C}_{lg} = \left[e^{j\lambda_l} \vdots e^{j2\lambda_l} \vdots \cdots \vdots e^{jn_a\lambda_l} \vdots 0 \vdots \cdots \vdots 0\right]^T.$$

Theorem 4. Under (AS1)-(AS4) and undermodeling, $\sqrt{\frac{TM}{L_T}}\left(\widehat{\theta}^{(1)}_{TM} - \overline{\theta}^{(1)}\right)$ is asymptotically Gaussian with zero-mean and covariance matrix specified by (5-1). The same result holds true under sufficient order modeling if $\mathcal{D}^{(so)}$ is a singleton. •

TABLE I : σ_{th} – theoretical σ, σ_{sim} – σ from simulations

Parameters		a_1	a_2	b_1	b_2
True Values		−1.500	0.700	1.000	0.500
Approach	estimate statistics: T = 1024, SNR=10dB				
Proposed (Sec. 3.2) (freq-dom.)	mean	−1.465	0.668	0.991	0.524
	σ_{sim}	±0.017	±0.015	±0.044	±0.058
	σ_{th}	±0.014	±0.012	±0.040	±0.050
	RMS	(0.039)	(0.036)	(0.045)	(0.063)
Least-Sq. (time-dom.)	mean	−0.710	−0.040	0.992	0.807
	σ_{sim}	±0.035	±0.033	±0.045	±0.044
	RMS	(0.791)	(0.740)	(0.046)	(0.310)
[1] (time-dom.)	mean	−1.452	0.627	0.989	0.514
	σ_{sim}	±0.038	±0.038	±0.041	±0.081
	RMS	(0.061)	(0.082)	(0.042)	(0.082)

6 Simulation Example

The true system model is

$$H(q) = \frac{q^{-1} + 0.5q^{-2}}{1 - 1.5q^{-1} + 0.7q^{-2}}$$

and the measured input is $u(t) = (1 + 0.8q^{-1})^{-1}\, w(t)$ where $\{w(t)\}$ is i.i.d., binary (± 1 with prob. 0.5 each). The measurement noise is colored Gaussian given by $v(t) = (1 - 0.95q^{-1})\epsilon(t)$ where $\epsilon(t)$ is i.i.d. zero-mean Gaussian.

The simulation results based on 100 Monte carlo runs are shown in Table I for the approaches proposed in Sec. 3.2 under the sufficient-order case with $n_a = n_{a0} = 2$ and $n_b = n_{b0} = 2$. We also show the theoretical standard deviations for the parameter estimates. These were calculated using the expression (5-1) with mean values of the estimated parameters as $\overline{\theta}^{(1)}$. In applying the proposed approaches, we selected $2m_T + 1 = 11$ for the record length of $T = 1024$. The number of frequency points M was taken to be all the points on the coarse grid (3-7) that lie in $(0, \pi)$. For comparison we also show the results obtained using the classical least-squares algorithm ([4, Sec. 7.1]) and the modified least-squares algorithm of [1]. Both of these approaches are time-domain approaches. The approach of [1] is designed to provide unbiased estimates for model (1-1) when $\{v(t)\}$ is white. It should be noted that [1] does not fit a model such as (1-3); rather, [1] fits

$$\frac{\sum_{i=0}^{n} b_i q^{-i}}{1 + \sum_{i=1}^{n} a_i q^{-i}}. \qquad (6-1)$$

In Table I we do not display b_0.

7 References

[1] P.A. Regalia, "An unbiased equation error identifier and reduced-order approximations," *IEEE Trans. Signal Processing*, vol. SP-42, pp. 1397-1412, June 1994.

[2] P.A. Regalia and P. Stoica, "Stability of multivariable least-squares models," *IEEE Signal Processing Letters*, vol. 2, pp. 195-196, Oct. 1995.

[3] J.K. Tugnait and Y. Ye, "Stochastic system identification with noisy input-output measurements using polyspectra," *IEEE Trans. Automatic Control*, vol. AC-40, pp. 670-683, April 1995.

[4] T. Söderström and P. Stoica, *System Identification.* Prentice Hall Intern.: London, 1989.

[5] L. Ljung, *System Identification: Theory for the User.* Prentice-Hall: Englewood Cliffs, N.J., 1987.

[6] T. Soderstrom and P. Stoica, "On the stability of dynamic models obtained by least-squares identification," *IEEE Trans. Automatic Control*, vol. AC-26, pp. 575-577, April 1981.

[7] D.R. Brillinger, *Time Series Data Analysis and Theory.* Holt, Rhinehart & Winston,: New York, 1975.

[8] P.A. Regalia, *Adaptive IIR Filtering in Signal Processing and Control.* Marcel Dekkar: New York, 1995.

[9] C.T. Mullis and R.A. Roberts, "The use of second-order information in the approximation of discrete-time linear systems," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. ASSP-24, pp. 226-238, June 1976.

[10] J.K. Tugnait, "Stochastic system identification with noisy input using cumulant statistics," *IEEE Transactions on Automatic Control*, vol. AC-37, pp. 476-485, April 1992.