FREQUENCY DOMAIN ANALYSIS OF ADAPTIVE TRACKING ALGORITHMS

Brett Ninness *,[1] **Juan Carlos Gómez** **,[1]
Steven Weller ***

* *Department of Electrical and Computer Engineering, University of Newcastle, Australia.*
** *Department of Electrical and Computer Engineering, University of Newcastle, Australia.*
*** *Department of Electrical and Electronic Engineering, University of Melbourne, Australia.*

Abstract: In this paper, an analysis of the tracking performance of several adaptive algorithms is carried out for the case of model structures with fixed pole positions. Such structures have recently been proposed as an efficient generalisation of the common FIR model structure. The focus of this work is to analyse the tradeoff between noise sensitivity and tracking ability in the frequency domain by illustrating how it is influenced by such things as input and noise spectral densities, step size and, what is the emphasis of this paper, the choice of the fixed pole locations.

Keywords: Adaptive Algorithms, Error Analysis, Asymptotic Analysis, Frequency Domains,

1. INTRODUCTION

This paper is inspired by recent results (Guo and Ljung, 1995) that quantify the parameter space performance of adaptive algorithms, by work suggesting the utility of analysis of adaptive algorithms in the frequency domain (Egardt *et al.*, 1992; C. R. Johnson Jr., 1995), by work suggesting such analysis be simplified by considering high model orders (Gunnarsson and Ljung, 1989), and by recent work suggesting novel model structures for adaptive algorithms (Williamson and Zimmermann, 1996).

These model structures are generalisations of the popular FIR structure, but are more flexible in that the poles in the model structure need not all be fixed at the origin. As pointed out in (Williamson and Zimmermann, 1996), exploiting this flexibility can accrue many advantages in terms of estimation accuracy, while still retaining the desirable convergence properties enjoyed by adaptive FIR schemes.

Following the suggestions in (Egardt *et al.*, 1992; C. R. Johnson Jr., 1995), this paper provides a frequency domain analysis of the performance of these 'generalized FIR' methods in order to make explicit how the tradeoff between noise sensitivity and tracking ability is influenced by input and noise spectral densities, choice of step size, and (what is a main focus of this paper) the choice of fixed pole position. There are close relations between this work and those of Gunnarsson and Ljung (Gunnarsson and Ljung, 1989) who studied adaptive FIR algorithms in the frequency domain. Specifically (Gunnarsson and Ljung, 1989) provide the main idea of this paper which is to simplify error expressions by considering large model order.

A key tool used in this paper is to re–parameterize the fixed denominator model structure into an orthonormal form studied in (Ninness and Gustafsson, 1997) in order to facilitate the theoretical analysis. This strategy illustrates that an orthonormal parameterisation is an intrinsic part of estimation using *any* fixed denominator model structures. This arises since for recursive least squares (RLS) and Kalman filtering algorithms, the tracking error in the frequency domain can be quantified in terms of the orthonormal form whether or not the model structure is originally cast in this form. The orthonormal form makes explicit how the error depends on the choice of fixed poles.

[1] This work was supported by the Australian Research Council and the Centre for Integrated Dynamics and Control.

2. PROBLEM FORMULATION

This paper considers situations where an observed input sequence $\{u_k\}$ is related to an observed output sequence $\{y_k\}$ according to

$$y_k = G_k(q)u_k + \nu_k \tag{1}$$

where $\{\nu_k\}$ is a zero mean white noise process with variance $\mathbf{E}\{\nu_k^2\} = \sigma_\nu^2 < \infty$ and

$$G_k(q) = \sum_{n=0}^{\infty} g_k(n)q^{-n}$$

is a possibly time varying linear system with impulse response $\{g_k(n)\} \in \ell_2$. It is assumed that $\{u_k\}$ is a realisation of a stationary stochastic process with covariance function $R_u(\tau) = \mathbf{E}\{u_k u_{k-\tau}\}$ and associated spectral density $\Phi_u(\omega) = \sum_{\tau=-\infty}^{\infty} R_u(\tau)e^{-j\omega\tau}$ and that $\{u_k\}$ is weakly uncorrelated with $\{\nu_k\}$ in the sense that $|\mathbf{E}\{u_k\nu_{k-\tau}\}| \to 0$ as $\tau \to \infty$. It is also assumed that $\Phi_u(\omega) > 0$ and that $\Phi_u(\omega)$ has a finite dimensional spectral factorisation.

At issue is the estimation of the (assumed unknown) time varying dynamics $G_k(q)$ by means of the observations $\{u_k\}$ and $\{y_k\}$. There are many approaches to this problem, but a common theme (Goodwin and Sin, 1984) is to express the dependence (1) in a linear regression form $y_k = \phi_k^T\theta_k + \nu_k$ where the 'regression vector' ϕ_k depends on measurements of $\{u_t\}$ and $\{y_t\}$ up until $t = k$ and $\theta_k \in \mathbf{R}^p$ is a vector of p parameters in a model structure $G(q, \theta_k)$ that attempts to describe the true dynamics $G_k(q)$. An estimate of $G_k(q)$ is then obtained as $G(q, \widehat{\theta}_k)$ where the estimate $\widehat{\theta}_k$ is obtained recursively via

$$\widehat{\theta}_{k+1} = \widehat{\theta}_k + L_k(y_k - \phi_k^T\widehat{\theta}_k), \quad \mu \in (0,1) \tag{2}$$

where L_k is a gain vector that may be computed in various ways. A common choice for this gain vector is $L_k = \mu\phi_k$, $\mu \in (0,1)$ in which case (2) is known as the 'gradient' or 'least mean square' (LMS) algorithm. Another common choice is $L_k = P_k\phi_k$ where P_k satisfies

$$P_k = \frac{1}{\lambda}\left\{P_{k-1} - \frac{P_{k-1}\phi_k\phi_k^T P_{k-1}}{\lambda + \phi_k^T P_{k-1}\phi_k}\right\}$$

with $\lambda = 1 - \mu, \mu \in (0,1)$ and P_k is initialised with some positive definite P_0 and with the ensuing algorithm being known as 'Recursive Least Squares' (RLS). Finally, if the time variation of the parameters θ_k are modeled via a random walk as $\theta_{k+1} = \theta_k + \rho\, w_k$ where w_k is a stationary zero mean vector white noise process with $\mathbf{E}\{w_k w_k^T\} = Q$, then the update law

$$L_k = \frac{\mu P_{k-1}\phi_k}{\sigma^2 + \mu\phi_k^T P_{k-1}\phi_k} \tag{3}$$

where P_k satisfies the Riccati equation

$$P_k = P_{k-1} - \mu\frac{P_{k-1}\phi_k\phi_k^T P_{k-1}}{\sigma^2 + \mu\phi_k^T P_{k-1}\phi_k} + \mu\Sigma \tag{4}$$

with $\Sigma > 0$ and symmetric is known as the Kalman Filter.

When employing any of these adaptive schemes, a central question is the accuracy of the estimate $G(q, \widehat{\theta}_k)$ as a description of $G_k(q)$. The most common way of assessing this is to examine the accuracy of $\widehat{\theta}_k$ itself (Goodwin and Sin, 1984). This may be achieved by defining θ_k as the true parameter vector that allows the model structure to exactly describe the underlying time varying dynamics as $G(q, \theta_k) = G_k(q)$ and by defining the estimation error $\widetilde{\theta}_k$ as $\widetilde{\theta}_k \triangleq \theta_k - \widehat{\theta}_k$.

Substituting this definition into the general update equation (2) gives that this error satisfies the following difference equation

$$\widetilde{\theta}_{k+1} = \left(I - L_k\phi_k^T\right)\widetilde{\theta}_k + \rho\,\omega_k - L_k\nu_k. \tag{5}$$

The quality of an adaptive estimation scheme can then be quantified by using (5) to calculate the covariance $\mathbf{E}\{\widetilde{\theta}_k\widetilde{\theta}_k^T\}$ as a measure of estimation accuracy. Unfortunately, as pointed out in (Gunnarsson and Ljung, 1989; Guo and Ljung, 1995), the exact expression for this covariance will be very complicated except in very special circumstances. The main result of (Guo and Ljung, 1995) which will be central to the analysis of this paper is that under the stated assumptions, $\mathbf{E}\{\widetilde{\theta}_k\widetilde{\theta}_k^T\}$ may be approximated by Π_k given by the deterministic difference equation

$$\begin{aligned}\Pi_{k+1} = (I - \mu S_k R)\Pi_k(I - \mu S_k R)^T + \\ \mu^2\sigma_\nu^2 S_k R S_k + \rho^2 Q\end{aligned} \tag{6}$$

where $R \triangleq \mathbf{E}\{\phi_k\phi_k^T\}$ and S_k is defined as

LMS:

$$S_k = I, \tag{7}$$

RLS:

$$S_k = (1+\mu)S_{k-1} - \mu S_{k-1}RS_{k-1}; S_0 = P_0, \tag{8}$$

Kalman Filter:

$$S_k = S_{k-1} - \mu S_{k-1}RS_{k-1} + \frac{\mu}{\sigma^2}\Sigma; S_0 = \frac{1}{\sigma^2}P_0, \tag{9}$$

In (Guo and Ljung, 1995) the quality of this approximation is quantified as $\left\|\mathbf{E}\left\{\widetilde{\theta}_k\widetilde{\theta}_k^T\right\} - \Pi_k\right\| \leq \kappa(\mu)$ where $\kappa(\mu)$ is a bounded function that tends to zero as μ tends to zero.

However, as argued in (Gunnarsson and Ljung, 1989; Egardt *et al.*, 1992), in many cases the interest is not in the accuracy in parameter space, but the accuracy in how close the estimated model $G(q, \widehat{\theta}_k)$ is to the true system $G_k(q)$ in terms of the error $\widetilde{G}_k(e^{j\omega}) \triangleq G_k(e^{j\omega}) - G(e^{j\omega}, \widehat{\theta}_k)$ in the estimated frequency response. In this paper, model structures $G(q, \theta_k)$ are considered for which the estimated frequency response depends linearly on the estimated parameters as $G(e^{j\omega}, \widehat{\theta}_k) = \Gamma_p^T(e^{j\omega})\widehat{\theta}_k$ where

$$\Gamma_p(q) \triangleq [\mathcal{B}_0(q), \mathcal{B}_1(q), \cdots, \mathcal{B}_{p-1}(q)]^T \quad (10)$$

is a vector of p rational transfer functions $\mathcal{B}_n(q)$. For example, $\mathcal{B}_n(q) = q^{-n}$ corresponds to an FIR model structure.

Using (6) and (10), an approximate frequency domain quantification of adaptive performance may then be taken as

$$\mathbf{E}\left\{|\widetilde{G}_k(e^{j\omega})|^2\right\} = \Gamma_p^\star(e^{j\omega})\mathbf{E}\left\{\widetilde{\theta}_k\widetilde{\theta}_k^T\right\}\Gamma_p(e^{j\omega})$$
$$\approx \Gamma_p^\star(e^{j\omega})\Pi_k\Gamma_p(e^{j\omega}) \quad (11)$$

where $\cdot^\star$ denotes 'conjugate transpose'. Unfortunately, again this expression will in general be of a very complicated nature. The main contribution of this paper will be to follow the lead of (Gunnarsson and Ljung, 1989) and derive simple approximations for (11) that are increasingly accurate for increasing model order p. These simplified expressions make clear how factors such as measurement noise variance, input spectral density, and (what is the novel part of this work) choice of fixed pole location affect $\mathbf{E}\{|\widetilde{G}_k(e^{j\omega})|^2\}$.

3. MODEL STRUCTURES

The model structures examined in this paper have recently been proposed and examined in an adaptive filtering context by Williamson and co–workers in a series of papers (see (Williamson and Zimmermann, 1996) for references) where they have been termed 'fixed pole adaptive filters'. They are formulated as

$$G(q, \theta'_k) = \left[\prod_{n=0}^{p-1}(q - \xi_n)\right]^{-1}\sum_{n=0}^{p-1}\theta'_k(n)q^n \quad (12)$$

where the poles $\{\xi_n\}$ are fixed according to prior information about the likely pole positions of the true time varying system $G_k(q)$. A special case of this structure arises when all the poles $\{\xi_n\}$ are chosen at the origin in which case (12) is an FIR model structure.

However, empirical evidence (Williamson and Zimmermann, 1996) supports the fact that in an adaptive filtering context, a significant improvement in estimation accuracy is possible by avoiding poles all fixed at the origin, and instead distributing them in the unit disk so as to be as close as possible to the true poles of $G_k(q)$.

In spite of the pleasant properties enjoyed by the model structure (12), its generality (as compared to an FIR structure) makes frequency domain analysis of adaptive algorithms much more difficult. To be more specific, with an FIR structure, the frequency response of the estimated model can be considered to be a linear combination of the 'basis functions' $\{1, e^{-j\omega}, \cdots, e^{-j(p-1)\omega}\}$ used in classical Fourier analysis.

Furthermore, since these FIR 'basis functions' enjoy the group structure $e^{-j\omega n}e^{-j\omega m} = e^{-j\omega(m+n)}$, then for large k the parameter covariance matrix approximation Π_k is Toeplitz, and by drawing on the wealth of literature on such matrices (Grenander and Szegö, 1958) it is possible to determine the function for which (11) is a (partial) Cèsaro mean Fourier reconstruction. These are the main tools used in (Gunnarsson and Ljung, 1989).

Unfortunately, for the more general model structure (12), all these properties are lost, making the instructive frequency domain expressions presented in (Gunnarsson and Ljung, 1989) less straightforward to derive. The contribution of this paper is to overcome these difficulties by replacing the model structure (12) with the following formulation

$$G(q, \theta_k) = \sum_{n=0}^{p-1}\theta_k(n)\mathcal{B}_n(q) \quad (13)$$

where

$$\mathcal{B}_n(q) = \left(\frac{\sqrt{1 - |\xi_n|^2}}{q - \xi_n}\right)\prod_{k=0}^{n-1}\left(\frac{1 - \overline{\xi_k}q}{q - \xi_k}\right) \quad (14)$$

This model structure (13) is easily cast in linear regression form with

$$\phi_k \triangleq \Gamma_p^T(q)u_k = [\mathcal{B}_0(q)u_k, \mathcal{B}_1(q)u_k, \cdots, \mathcal{B}_{p-1}(q)u_k],$$
$$\theta_k \triangleq [\theta_k(0), \theta_k(1), \cdots, \theta_k(p-1)]^T.$$

A key point is that since the poles of the model structure (13) and (12) are identical, then they are equivalent in the sense that for some non-singular $J \in \mathbf{R}^{p\times p}$, the parameter vectors θ_k in (13) and $\theta'_k = J\theta_k$ in (12) describe exactly the same transfer function. As well, with initialisation $P_0 = J^{-1}P'_0J^{-T}$ consistent with this linear re–parameterisation, the RLS update equations are invariant to the re–parameterisation in the sense that $\widehat{\theta}'_k = J\widehat{\theta}_k$ so that frequency response estimates are identical: $G(e^{j\omega}, \widehat{\theta}'_k) = G(e^{j\omega}, \widehat{\theta}_k)$. This same property also applies to the Kalman Filtering update law (3),(4) provided the compatibility $\Sigma = J^{-1}\Sigma' J^{-1}$ is also maintained.

4. STEADY STATE ANALYSIS

The steady state behavior of the frequency response estimation error is defined to be the quantity $\mathbf{E}\{|\widetilde{G}_k(e^{j\omega})|^2\}$ for large k, and in order for it to be investigated, it is necessary to examine the behavior of the solution Π_k of (6) for large k; that is $\lim_{k\to\infty}\Pi_k \triangleq \Pi$ where Π may be evaluated by determining the steady state solutions $S \triangleq \lim_{k\to\infty} S_k$ of (7)–(9) and then substituting them into (6) before examining its steady own state solution with terms of order $\mu^2\Pi$ discarded (Guo and Ljung, 1995).

LMS: Here $S = I$ so that Π is the solution of the Lyapunov equation

$$\Pi R + R\Pi = \mu\sigma_\nu^2 R + \frac{\rho^2}{\mu}Q. \qquad (15)$$

RLS: Here $S = R^{-1}$ so that Π is given by

$$\Pi = \frac{\mu\sigma_\nu^2}{2}R^{-1} + \frac{\rho^2}{2\mu}Q. \qquad (16)$$

Kalman Filter: This case is more difficult. S is the solution of $\sigma^2 SRS = \Sigma$ so that Π is given by the solution of

$$SR\Pi + \Pi RS = \frac{\mu\sigma_\nu^2}{\sigma^2}\Sigma + \frac{\rho^2}{\mu}Q. \qquad (17)$$

For the special case of $\Sigma = Q$ this system has solution

$$\Pi = \frac{\sigma^2}{2}\left(\mu\frac{\sigma_\nu^2}{\sigma^2} + \frac{\rho^2}{\mu}\right)S. \qquad (18)$$

Using these characterisations of Π together with (11) it is possible to quantify the steady state estimation error in the frequency domain as

$$\mathbf{E}\left\{|\widetilde{G}(e^{j\omega})|^2\right\} \triangleq \lim_{k\to\infty}\mathbf{E}\left\{|\widetilde{G}_k(e^{j\omega})|^2\right\} \approx \Gamma_p^\star(e^{j\omega})\Pi\Gamma_p(e^{j\omega}). \qquad (19)$$

Although this provides a quantifiable performance measure, the resulting expression is so complicated that it is difficult to extract useful design insight from it. However, if the model order is assumed large, then the expression (19) can be significantly simplified, and doing so illustrates the crucial dependence of the error (19) on the term $\gamma_p(\omega)$ defined as

$$\gamma_p(\omega) \triangleq \sum_{k=0}^{p-1}|\mathcal{B}_k(e^{j\omega})|^2, \qquad (20)$$

the nature of which is illustrated in figure 1. In the results that follow, the quantity $\delta(\omega) \triangleq \mathbf{E}\left\{|G_{k+1}(e^{j\omega}) - G_k(e^{j\omega})|^2\right\}$ is used to quantify the time variation of $G_k(q)$ in the frequency domain.

Theorem 4.1. For the LMS algorithm and the model structure (13), then

$$\lim_{p\to\infty}\frac{1}{\gamma_p(\omega)}\mathbf{E}\left\{|\widetilde{G}(e^{j\omega})|^2\right\} = \frac{1}{2}\left[\mu\sigma_\nu^2 + \frac{\rho^2\delta(\omega)}{\mu\Phi_u(\omega)}\right]$$

Proof: Using Parseval's Theorem together with the notation (A.1) in (15) gives that in the limit as $k\to\infty$

$$\frac{\Gamma_p^\star(e^{j\omega})\Pi M_p(\Phi_u)\Gamma_p(e^{j\omega})}{\gamma_p(\omega)} + \frac{\Gamma_p^\star(e^{j\omega})M_p(\Phi_u)\Pi\Gamma_p(e^{j\omega})}{\gamma_p(\omega)} = \frac{\mu\sigma_\nu^2\Gamma_p^\star(e^{j\omega})M_p(\Phi_u)\Gamma_p(e^{j\omega})}{\gamma_p(\omega)} + \frac{\rho^2\delta_p(\omega)}{\mu\gamma_p(\omega)}.$$

Taking the limit of both sides as $p\to\infty$ while using Theorem A.1 then gives the result. ■

The interpretation of this theorem is that for large model order p, and after the algorithm has converged (large k)

$$\mathbf{E}\left\{|\widetilde{G}_k(e^{j\omega})|^2\right\} \approx \frac{\gamma_p(\omega)}{2}\left[\mu\sigma_\nu^2 + \frac{\rho^2\delta(\omega)}{\mu\Phi_u(\omega)}\right] \quad (21)$$

For the case of all the poles $\{\xi_n\}$ in the model structure $G(q,\theta)$ chosen at the origin, then $\gamma_p(\omega) = p$ and the above expression specialises to that derived in (Gunnarsson and Ljung, 1989). However, as illustrated in figure 1, the factor $\gamma_p(\omega)$ in the expression (21) shows how pole choices other than FIR influence the frequency domain estimation error.

Of course, a fundamental question is the reliability of using the approximation (21) for practically useful (and hence finite) model orders, given that it is obtained from a result that is asymptotic in p. The most suitable way to deal with this issue would be to quantify the convergence rate in theorem 4.1. This appears to be extremely difficult. Instead, the approach used in (Gunnarsson and Ljung, 1989) is taken wherein the validity of (21) for finite p is examined via a simulation study. This is done in §5, where it is shown in figure 2 that for a tenth order model, (21) is quite an accurate approximation; see (Ninness and Gómez, 1996) for an illustration of the veracity of the approximation for only a 4th order model structure.

Theorem 4.2. For the RLS algorithm and the model structure (12) or (13), then

$$\lim_{p\to\infty}\frac{1}{\gamma_p(\omega)}\mathbf{E}\left\{|\widetilde{G}(e^{j\omega})|^2\right\} = \frac{1}{2}\left[\frac{\mu\sigma_\nu^2}{\Phi_u(\omega)} + \frac{\rho^2}{\mu}\delta(\omega)\right]$$

Proof: Using Parseval's Theorem together with the notation (A.1) in (16) gives that

$$\frac{1}{\gamma_p(\omega)}\mathbf{E}\left\{|\widetilde{G}(e^{j\omega})|^2\right\} = \frac{\Gamma_p^\star(e^{j\omega})\Pi\Gamma_p(e^{j\omega})}{\gamma_p(\omega)} =$$
$$\frac{\mu\sigma_\nu^2}{2\gamma_p(\omega)}\Gamma_p^\star(e^{j\omega})M_p^{-1}(\Phi_u)\Gamma_p(e^{j\omega}) + \frac{\rho^2\delta_p(\omega)}{2\mu\gamma_p(\omega)}$$

Taking the limit of both sides as $p \to \infty$ while using Theorem A.1 then gives the result. ■

Following the example of the previous theorem, the interpretation of this theorem is that for large model order p, and after the algorithm has converged (large k)

$$\mathbf{E}\left\{|\widetilde{G}_k(e^{j\omega})|^2\right\} \approx \frac{\gamma_p(\omega)}{2}\left[\frac{\mu\sigma_\nu^2}{\Phi_u(\omega)} + \frac{\rho^2}{\mu}\delta(\omega)\right]. \quad (22)$$

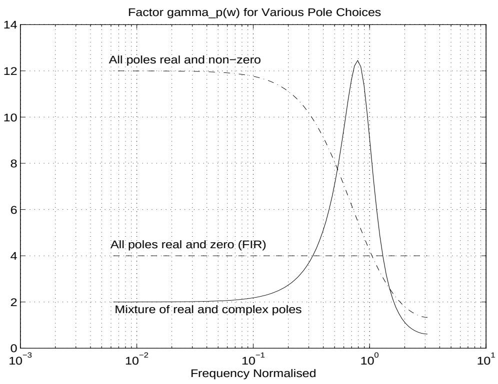

Fig. 1. *Plot of* $\gamma_p(\omega) = \sum_{k=0}^{p-1}|\mathcal{B}_k(e^{j\omega})|^2$ *for* $p = 4$ *and for various choices of pole location.*

To complete the analysis, the following theorem quantifies Kalman Filter performance However, there are particular difficulties in solving for the steady state parameter covariance Π, and this leads to the treatment of only a specialised case in which $Q = \Sigma$.

Theorem 4.3. For the Kalman Filtering algorithm, the model structure (12) or (13) and under the assumption that $Q = \Sigma$, then

$$\lim_{p\to\infty}\frac{1}{\gamma_p(\omega)}\mathbf{E}\left\{|\widetilde{G}(e^{j\omega})|^2\right\} = \frac{1}{2}\left(\mu\frac{\sigma_\nu^2}{\sigma^2} + \frac{\rho^2}{\mu}\right)\sqrt{\frac{\sigma^2\delta(\omega)}{\Phi_u(\omega)}}$$

Proof: Uses the same ideas as in the previous proofs, but is more arithmetically complicated. See (Ninness and Gómez, 1996) for details. ■

In sympathy with previous results, the interpretation of this theorem is that for large model order p and after the algorithm has converged (large k) then

$$\mathbf{E}\left\{|\widetilde{G}_k(e^{j\omega})|^2\right\} \approx \frac{\gamma_p(\omega)}{2}\left(\mu\frac{\sigma_\nu^2}{\sigma^2} + \frac{\rho^2}{\mu}\right)\sqrt{\frac{\sigma^2\delta(\omega)}{\Phi_u(\omega)}} (23)$$

The ubiquity of the term $\gamma_p(\omega)$ in all these error quantifications shows that the orthonormal parameterisation (13),(14) is more than just an essential tool for the analysis of general fixed denominator model structures. Instead, the orthonormal 'basis functions' $\{\mathcal{B}_n(q)\}$ appear as an intrinsic part of adaptive estimation with any fixed denominator model structure $G(q, \theta)$.

5. SIMULATION EXAMPLE

In this section, the utility of the previous theoretical analysis will be demonstrated via a simulation study in which it is supposed that there is an underlying system

$$G(q) = \frac{0.0355q + 0.0247}{(q - 0.9048)(q - 0.3679)} \quad (24)$$

from which input–output data is collected when the input $\{u_k\}$ is stationary and Gaussian with spectral density $\Phi_u(\omega) = 10/(1.25 - \cos\omega)$, and the observed output $\{y_k\}$ is subject to white Gaussian corruption $\{\nu_k\}$ of variance $\sigma_\nu^2 = 0.01$. Based on this observed data, an attempt is made to estimate $G(q)$ using the model structure (13) with poles $\{\xi_n\}$ chosen to correspond to continuous time guesses of 0.2 and 0.25 radians per second. Note that these poles, being far from either of the true poles at 0.1 and 1 rad/s, are particularly bad guesses. They have been chosen to dispel any suspicion in the sequel that the high accuracy of the approximations (21),(22) and (23) illustrated in figure 2 derives from unreasonable prior knowledge or idealised conditions.

All three algorithms, the LMS with $\mu = 0.001$, RLS with $\lambda = 0.999$ and $P_0 = I$, and the Kalman Filter with $\mu = 0.001, \Sigma = 0.1, P_0 = I$ and $\sigma^2 = 0.01$ were employed with a tenth order model structure ($p = 10$).

These estimation experiments were performed five hundred times with different realisations for the input and measurement noise. This allowed the true frequency domain estimation error $\mathbf{E}\{|\widetilde{G}_k(e^{j\omega})|^2\}$ to be estimated by calculating its sample value as an average over the 500 realisations. This is plotted as the solid line in figure 2. The dash–dot lines are the approximations (21), (22) and (23) derived from theorems 4.1–4.3.

In all cases, the close agreement between the solid and dashed plots shows the approximations (21), (22) and (23) to be highly accurate. This is in spite of the approximations being derived from asymptotic in p results, but being applied in this simulation to only a $p = 10$'th order model.

6. CONCLUSIONS

This paper has provided an analysis of the frequency domain error for various adaptive estimation schemes. This results in the extension of certain results already known for FIR model structures wherein all poles are fixed at the origin, to more general model structures where the poles may be placed arbitrarily, so long as they are stable. It was shown that the entire effect of the choice of fixed pole position on the frequency domain steady–state error depends on one term (called $\gamma_p(\omega)$). For FIR models, this term is simply p, the model order, but for fixed pole choices not at the origin (FIR), the term becomes frequency dependent in a manner that is influenced solely by the choice of poles.

Incidentally, it would be reasonable to suspect that all the results presented in this paper could be derived by a much simpler strategy of employing known results for FIR model structures (Gunnarsson and Ljung, 1989) with the input spectrum considered to be $\Phi_u(\omega)/|D_p(e^{j\omega})|^2$ where $D_p(q) = \prod_{k=0}^{p-1}(q-\xi_k)$ is the fixed denominator being employed.

Unfortunately, this simple approach fails since it leads to the conclusion that the error is invariant to the choice of pole location, and this can easily be seen to be incorrect via simple simulation. The failure of the method can be traced to a delicacy in which one is implicitly trying to recover a function $(\Phi_u(\omega)/|D_p(e^{j\omega})|^2)$ from it's partial Fourier series of length p, and since the function changes as p grows, the Fourier series does not converge with increasing p.

This paper sidesteps this difficulty by a strategy of re–parameterization with respect to a particular orthonormal basis under which the (generalized) Fourier series one is considering concerns a function $\Phi_u(\omega)$ which is invariant to p. The difficulty with this latter method, is that it requires certain new results (Theorem A.1) which generalise Toeplitz matrix and Cèsaro mean convergence results from the conventional trigonometric basis setting.

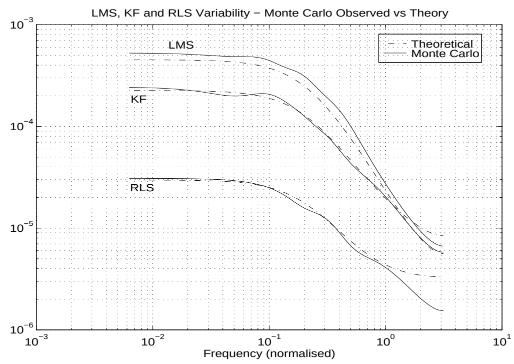

Fig. 2. *Comparison of observed variability (solid line) vs theory (dash–dot line) for Kalman Filter (top), LMS (middle) and RLS (bottom).*

Appendix A. TECHNICAL RESULTS

Define, for a function $f(\omega) > 0$ the matrix

$$M_p(f) \triangleq \frac{1}{2\pi}\int_{-\pi}^{\pi} \Gamma_p(\sigma)\Gamma_p^\star(\sigma)f(\sigma)\,\mathrm{d}\sigma. \quad \text{(A.1)}$$

Theorem A.1. Suppose $f(\omega)$ is continuous. Then provided $\sum_{k=0}^{\infty}(1-|\xi_k|) = \infty$

$$\lim_{p\to\infty} \frac{\Gamma_p^\star(\omega)M_p(f)\Gamma_p(\omega)}{\gamma_p(\omega)} = f(\omega).$$

If in addition f has a finite dimensional spectral factorisation then

$$\lim_{p\to\infty} \frac{\Gamma_p(\mu)^\star M_p^{-1}(f)\Gamma_p(\omega)}{\gamma_p(\omega)} = f^{-1}(\omega).$$

Additionally, if $Q_p \in \mathbf{R}^{p\times p}$ is symmetric, positive definite and $\|Q_p\|_2 < \infty$ for all p, then

$$\lim_{p\to\infty} \frac{1}{\gamma_p(\omega)}\Gamma_p^\star(\omega)M_p(f)Q_pM_p(f)\Gamma_p(\omega) =$$

$$f^2(\omega)\lim_{p\to\infty}\frac{1}{\gamma_p(\omega)}\Gamma_p^\star(\omega)Q_p\Gamma_p(\omega).$$

$$\lim_{p\to\infty} \frac{\Gamma_p^\star(\omega)M_p(f)Q_p\Gamma_p(\omega)}{\gamma_p(\omega)} = f(\omega)\lim_{p\to\infty}\frac{\Gamma_p^\star(\omega)Q_p\Gamma_p(\omega)}{\gamma_p(\omega)}.$$

Proof: See (Ninness and Gómez, 1996). ■

Appendix B. REFERENCES

C. R. Johnson Jr. (1995). On the interaction of adaptive filtering, identification and control. *IEEE Signal Processing Magazine.*

Egardt, B., C. R. Johnson Jr., L. Ljung and G.Ã. Williamson (1992). Adaptive system performance in the frequency domain. In: *Proc. 4th IFAC Symp. Adapt. Sys. Control and Sig. Proc.*. pp. 33–39.

Goodwin, G.C. and K.W. Sin (1984). *Adaptive Filtering Prediction and Control.*

Grenander, U. and G. Szegö (1958). *Toeplitz Forms and their Applications.*

Gunnarsson, S. and L. Ljung (1989). Frequency domain tracking characteristics of adaptive algorithms. *IEEE TASSP* **37**(7), 1072–1089.

Guo, Lei and Lennart Ljung (1995). Performance analysis of general tracking algorithms. *IEEE Trans. AC* **40**(8), 1388–1402.

Ninness, B. and F. Gustafsson (1997). A unifying construction of orthonormal bases for system identification. *IEEE Trans. AC* **42**(4).

Ninness, B. and J.C. Gómez (1996). Tech. Rep. EE9656. Dept. EE & CE, Uni. Newc. Obtainable via `http://www.ee.newcastle.edu.au`.

Williamson, G.A. and S. Zimmermann (1996). Globally convergent adaptive IIR filters based on fixed pole locations. *IEEE Trans SP* **44**(6), 1418–1427.