

Parameter Estimation Of A Continuous-Time Plant – The Least-Absolute Error With Variable Forgetting Factor Method

Hazlina Selamat¹, Rubiyah Yusof¹, Roger M. Goodall²

¹Electrical Eng. Faculty, Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, MALAYSIA

²Electrical & Electronics Eng. Dpt., Loughborough University, Leicestershire LE11 3TU, UNITED KINGDOM
hazlina@fke.utm.my

Abstract

The least-absolute error with variable forgetting factor (LAE+VFF) estimation method is proposed in this paper to estimate the parameters of time-varying continuous-time (C-T) systems. The algorithm includes an instrumental variable (IV) element to reduce estimation bias and a variable forgetting factor for good parameter tracking and smooth steady state.

1. Introduction

Advances in digital computers have contributed to the popularity of discrete-time (D-T) approaches in the area of system identification. However, real-world systems are continuous in nature and a D-T model does not have direct physical interpretation of a system compared to a C-T model. There are also problems associated with control and identification of C-T systems in D-T such as those described in [1]. Therefore, parameter estimation based on C-T system models have been developed. Most of them, however, deal only with open-loop systems. When data for parameter estimation process are obtained from systems operating in closed loop, such as in feedback control systems, the estimates are biased because the input and the noise signal (measurement or process noise) are correlated due to feedback [2]. To overcome bias problem, method such as in [3] has been suggested. This method however, requires measurement of additional excitation signal, which may not available in the system under consideration. If this extra signal is introduced, its effects on the overall system may not be desirable. Another approach is to use an IV technique [4], which is also effective in overcoming bias problems. The least absolute error (LAE) method has an IV-like element, inherent in the algorithm that can reduce bias in the estimates. It is originally designed for time-invariant systems but the inclusion of variable forgetting factor (VFF) in this paper allows it to deal with systems experiencing infrequent abrupt parameter changes.

2. System description

A C-T model described by Eq. (1) is considered.

$$G(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0} = \frac{Y(s)}{U(s)} \quad (1)$$

s is the differential operator, d/dt , and $n > m \geq 0$. In the C-T domain, Eq. (1) can be written as

$$y_n(t) = [-y_{n-1}(t) \ \dots \ -y(t) \ u_m(t) \ \dots \ u(t)] \times [a_{n-1} \ \dots \ a_0 \ b_m \ \dots \ b_0]^T \quad (2)$$

$$\text{where } y_i(t) = \frac{d^i y(t)}{dt^i} \text{ and } u_i(t) = \frac{d^i u(t)}{dt^i}$$

Including $\eta(t)$, which represents the measurement noise affecting the output, Eq. (2) can then be written in the following regression model form:

$$y_{diff}(t) = \phi_{diff}^T(t) \theta + \eta(t) \quad (3)$$

where

$$y_{diff}(t) = y_n(t)$$

$$\phi_{diff}^T(t) = [-y_{n-1}(t) \ \dots \ -y(t) \ u_m(t) \ \dots \ u(t)]$$

$$\theta = [a_{n-1} \ \dots \ a_0 \ b_m \ \dots \ b_0]^T$$

To estimate θ that contains parameters to be identified, the derivatives of the input and output signals are needed, but differentiating the possibly noisy signals would only accentuate the noise. To avoid direct differentiation of noisy data, the dynamical system described by differential equations (Eq. (2)), has to be converted into a system described by algebraic equations. This approach is known as ‘linear dynamic operation’ [5], where sampled input and output signals are passed through the linear dynamic operator (\mathcal{LD}) and the operator’s outputs can be used directly in the estimation algorithm as for D-T systems (Fig. 1).

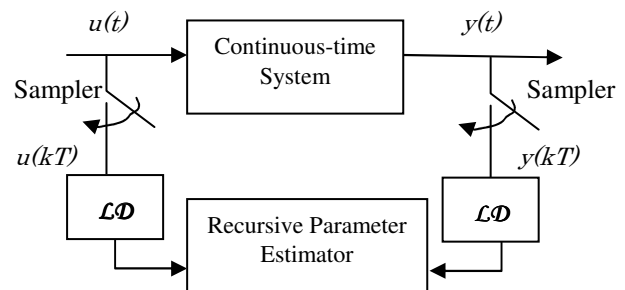


Fig. 1: Linear dynamic operation for C-T parameter estimation

This is a ‘direct’ identification method of C-T systems and does not involve any conversion of continuous- to discrete-time model or vice-versa. One manifestation of LD operation to handle time-derivative

terms in the regression vector, $\phi(t)$, is by using the Linear Integral Filter (LIF) [6], which is a digital implementation of multiple integral operations via numerical integration such as trapezoidal rule, rectangular (Euler's) approximation and parabolic (Simpson's) rule.

Firstly, $\phi(t)$ containing multiple integrals instead of the derivatives of the input and output signals are used. To do this, a finite horizon n^{th} - order integration operator, Γ^n , is introduced and multiple integrations over the time interval $[t-l_f T, t)$ are performed on both sides of Eq. (3), giving Eq. (4). T is the step size (taken the same as the sampling interval, T_s) and l_f is called a length factor of the LIF (a natural number).

$$y_\Gamma(t) = \phi_\Gamma^T(t)\theta + \eta_\Gamma(t) \quad (4)$$

where,

$$\begin{aligned} y_\Gamma(t) &= \Gamma^n y_n(t) \\ \phi_\Gamma^T(t) &= [-\Gamma^n y_{n-1}(t) \quad \dots \quad -\Gamma^n y(t) \quad \Gamma^n u_m(t) \quad \dots \quad \Gamma^n u(t)] \\ \theta(t) &= [a_{n-1} \quad \dots \quad a_0 \quad b_m \quad \dots \quad b_0]^T \end{aligned}$$

The C-T multiple integrals in the regression vector, $\phi(t)$, of Eq. (4) can be replaced with their D-T equivalent using 'normal method of numerical integration utilizing spline-based interpolation of sampled input and output data' [7], which is actually the generalization of the LIF method. The proposed LAE+VFF algorithm will then be derived, and is based on the regression model of the system, given by Eq. (4).

3. Least-absolute error with variable

forgetting factor (LAE+VFF) estimation

3.1 Continuous-time LAE+VFF The C-T LAE + VFF estimation method minimises the criterion given by Eq. (5)

$$J_{LAE}(\theta) = \int_0^t \mu^{t-\tau} |\varepsilon(\tau)| d\tau \quad (5)$$

where $\varepsilon(\tau) = y_\Gamma(\tau) - \phi_\Gamma^T(\tau)\theta$, t and τ are normalised quantities of the same independent time variable and μ is the C-T VFF that has a value between 0 and 1. Minimisation of Eq. (5) with respect to θ gives

$$\frac{\partial J_{LAE}(\theta)}{\partial \theta} = \int_0^t \mu^{t-\tau} \phi_\Gamma(\tau) \frac{\varepsilon(\tau)}{|\varepsilon(\tau)|} d\tau = 0 \quad (6)$$

where $\varepsilon(\tau) = y_\Gamma(\tau) - \phi_\Gamma^T(\tau)\theta$. Since $\varepsilon(t)$ is not known *a priori*, it is assumed that the current estimate of $\varepsilon(t)$, $\tilde{\varepsilon}(t)$, is available, say from another predictor. Therefore, the approximate C-T LAE solution of Eq. (6) can be derived as

$$\hat{\theta}(t) = M^{-1}(t)N(t) \quad (7)$$

$$M(t) = \int_0^t \mu^{t-\tau} \frac{\phi_\Gamma(\tau)\phi_\Gamma^T(\tau)}{|\tilde{\varepsilon}(\tau)|} d\tau \quad (8)$$

$$N(t) = \int_0^t \mu^{t-\tau} \frac{\phi_\Gamma(\tau)y_\Gamma(\tau)}{|\tilde{\varepsilon}(\tau)|} d\tau \quad (9)$$

Letting $M^{-1}(t) = \tilde{P}(t)$ = parameter covariance matrix,

$$\dot{\tilde{P}}(t) = \frac{d}{dt} M^{-1}(t) = -\tilde{P}(t) \ln \mu - \frac{\tilde{P}(t)\phi_\Gamma(t)\phi_\Gamma^T(t)\tilde{P}(t)}{|\tilde{\varepsilon}(t)|} \quad (10)$$

The time derivative of $\hat{\theta}$ in Eq. (7) can be written as

$$\dot{\hat{\theta}}(t) = \frac{d}{dt} [\tilde{P}(t)N(t)] = \frac{\tilde{P}(t)\phi_\Gamma(t)\varepsilon(t)}{|\tilde{\varepsilon}(t)|} \quad (11)$$

Therefore, the C-T LAE estimator is as follows:

(i) Model error is evaluated

$$\varepsilon(t) = y(t) - \phi^T(t)\hat{\theta}(t) \quad (12)$$

(ii) Parameter covariance matrix is updated (from Eq. (10))

$$d\tilde{P}(t) = -\left[\tilde{P}(t) \ln \mu + \tilde{P}(t) \frac{\phi(t)}{|\tilde{\varepsilon}(t)|} \phi^T(t) \tilde{P}(t) \right] dt \quad (13)$$

(iii) The estimates are corrected (from Eq. (11))

$$d\hat{\theta}(t) = \tilde{P}(t) \frac{\phi(t)}{|\tilde{\varepsilon}(t)|} \varepsilon(t) dt \quad (14)$$

3.2 Discretisation of the regression vector An n^{th} -order multiple integration, Γ^n , of a C-T signal, $x(t)$, over the time interval $[t-l_f T, t)$ yields

$$\Gamma^n x(t) = \frac{(1 - e^{-sl_f T})^n}{s^n} x(t) = \int_{t-l_f T}^t \dots \int_{t_{n-1}-l_f T}^{t_{n-1}} x(t_n) dt_n \dots dt_1$$

Thus, the elements of $\phi_\Gamma(t)$ can be evaluated using an operator polynomial, $J_i^n(q^{-1})$ [7] as Eq. (15).

$$\Gamma^n x_i(t) = \frac{(1 - e^{-sl_f T})^n}{s^n} \cdot \frac{d^i x(t)}{dt^i} = J_i^n(q^{-1})x(kT) \quad (15)$$

where

$$\begin{aligned} J_i^n(q^{-1}) &= \frac{a!T^{n-i}}{(a+n-i)!} \cdot [D_{l_f}(q^{-1})]^{n-i} \cdot N_{a+n-i}(q^{-1}) \\ &\quad \times N_b(q^{-1}) \cdot (1 - q^{-l_f})^i \end{aligned}$$

$$D_{l_f}(q^{-1}) = 1 + q^{-1} + \dots + q^{-(l_f-1)} = \frac{1 - q^{-l_f}}{1 - q^{-1}}$$

q^{-1} = unit - delay operator

l_f = length of the discrete - time integration window

$i = 0, 1, \dots, n$

n = system order

a, b = orders of the input and output signal interpolation

$N_0(q^{-1}) = N_1(q^{-1}) \equiv 1$

$$N_b(q^{-1}) = \sum_{i=0}^{b-1} w_{b,b-i} q^{-i}$$

= certain 'normal' polynomial of degree (b - 1)

$$w_{b,1} = w_{b,b} = 1$$

$$w_{b,j} = (b+1-j)w_{b-1,j-1} + jw_{b-1,j}$$

Replacing the time index kT in Eq. (15) with k only and omitting the subscript Γ associated with the multiple integral of the output signal and the regression vector, the D-T regression model of Eq. (4) can be written as

$$y(k) = \phi^T(k)\theta + \eta(k) \quad (16)$$

where,

$$y(k) = J_n^n(q^{-1})y(k)$$

$$\phi^T(k) = \begin{bmatrix} -J_{n-1}^n(q^{-1})y(k) & \cdots & -J_0^n(q^{-1})y(k) \\ \cdots & J_m^n(q^{-1})u(k) & \cdots & J_0^n(q^{-1})u(k) \end{bmatrix}$$

T is generally selected according to

$$\frac{\pi}{150\omega_n} < T < \frac{\pi}{10\omega_n}, \text{ where } \omega_n \text{ is the system's natural}$$

frequency. l_f is chosen so that the frequency response of the LIF matches as closely as possible the frequency band of the system to be identified.

3.3 Discretisation of the estimator Rectangular approximation is then used to discretise the C-T LAE + VFF estimator given by Eq. (12) to Eq. (14) and the D-T regression model in Eq. (16) can be directly utilised. The rectangular integration scheme is applied to Eq. (8) to Eq. (9) to give

$$M(k) \approx T \sum_{i=1}^k \mu^{(k-i)T} \xi(i) \phi^T(i) = \gamma M(k-1) + T \xi(k) \phi^T(k) \quad (17)$$

$$N(k) \approx T \sum_{i=1}^k \mu^{(k-i)T} \xi(i) y(i) = \gamma N(k-1) + T \xi(k) y(k) \quad (18)$$

$$\hat{\theta}(k) = \hat{\theta}(k-1) + TM^{-1}(k) \xi(k) [y(k) - \phi^T(k) \hat{\theta}(k-1)] \quad (19)$$

where $\xi(k) = \frac{\phi(k)}{|\tilde{\varepsilon}(k)|}$ and $\gamma = \mu^T$ i.e the T^{th} -power of

the continuous-time VFF, μ , representing its discrete-time equivalent. Applying Matrix Inversion Lemma (MIL) to Equation (17), $M^{-1}(t) = \tilde{P}(t)$ can be evaluated as Eq. (20).

$$\tilde{P}(k) = \frac{1}{\gamma} \left[\tilde{P}(k-1) - T \frac{\tilde{P}(k-1) \xi(k) \phi^T(k) \tilde{P}(k-1)}{\gamma + T \phi^T(k) \tilde{P}(k-1) \xi(k)} \right] \quad (20)$$

Then, taking $P(k) = T \tilde{P}(k)$ = parameter covariance matrix scaled by the sampling period, the recursive D-T LAE can be summarized as follows:

$$\begin{aligned} \hat{\theta}(k) &= \hat{\theta}(k-1) + P(k) \xi(k) \varepsilon(k) \\ &= \hat{\theta}(k-1) + \frac{P(k-1)}{\gamma + \phi(k) P(k-1) \xi(k)} \xi(k) \varepsilon(k) \end{aligned} \quad (21)$$

$$P(k) = \frac{1}{\gamma} \left[P(k-1) - \frac{P(k-1) \xi(k) \phi^T(k) P(k-1)}{\gamma + \phi^T(k) P(k-1) \xi(k)} \right] \quad (22)$$

Eq. (21) and Eq. (22) show that the discretised LAE estimator has a similar recursive representation as in the discrete-time framework of IV estimation scheme known from discrete-time identification literature (e.g. [4]), which has the general form

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \Omega(k) \xi(k) \varepsilon(k) \quad (23)$$

where

$$\varepsilon(k) = y(k) - \phi^T(k) \hat{\theta}(k-1)$$

$$\xi(t) = \frac{\phi(t)}{|\tilde{\varepsilon}(t)|} = \frac{\phi(t)}{|y(k) - \phi^T(k) \hat{\theta}(k)|} = \text{IV vector}$$

$$\Omega(k) = \frac{P(k)}{\gamma + \phi^T(k) P(k) \xi(k)} = \text{gain matrix}$$

The on-line evaluation of the *a posteriori* prediction error, $\tilde{\varepsilon}(k)$, can be expressed as

$$\tilde{\varepsilon}(k) = \varepsilon(k) - \frac{1}{\gamma} \phi^T(k) P(k-1) \phi(k) \text{sgn}[\varepsilon(k)] \quad (24)$$

To track time-varying parameters, the choice of Ω in Eq. (23) is crucial. Large Ω results in good parameter tracking but the estimator becomes more sensitive to noise. In contrast, small Ω makes it less sensitive to noise, leading to smooth steady-state performance but giving poor parameter tracking. Thus, a variable gain is needed. Employing a VFF method in [8], the following RLAE + VFF algorithm is proposed:

(i) Calculate the *a priori* prediction error

$$\varepsilon(k) = y(k) - \phi^T(k) \hat{\theta}(k-1) \quad (25a)$$

(ii) Calculate the *a posteriori prediction error*

$$\tilde{\varepsilon}(k) = \varepsilon(k) - \frac{1}{\gamma(k)} \phi^T(k) P(k-1) \text{sgn}(\varepsilon(k)) \quad (25b)$$

(iii) Evaluate the IV vector,

$$\xi(k) = \begin{cases} \frac{\phi(k)}{|\tilde{\varepsilon}(k)|} & \text{if } |\tilde{\varepsilon}(k)| \geq \varepsilon_{\min} \\ 0 & \text{if } |\tilde{\varepsilon}(k)| < \varepsilon_{\min} \end{cases} \quad (25c)$$

(iv) Generate the new estimate

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \frac{P(k-1)}{\gamma(k) + \phi(k) P(k-1) \xi(k)} \xi(k) \varepsilon(k) \quad (25d)$$

(v) Update the parameter covariance matrix

$$P(k) = \frac{1}{\gamma(k)} \left[P(k-1) - \frac{P(k-1) \xi(k) \phi^T(k) P(k-1)}{\gamma(k) + \phi^T(k) P(k-1) \xi(k)} \right] \quad (25e)$$

(vi) Decide the value of γ which is the D-T equivalent of the C-T VFF, μ

$$\mu = \begin{cases} \mu_1 & \text{if } |\varepsilon(k)| > \kappa \\ \mu_2 & \text{if } |\varepsilon(k)| \leq \kappa \end{cases} \quad (25f)$$

$$\gamma(k) = \mu^T$$

κ is a design variable that balances the tracking rate and steady-state smoothness of the parameter estimates. By running the estimator with known, fixed system parameters several times, the choice of κ can be based on the estimation error. For example it can be chosen to be greater than the average estimation error so that when the error is greater than the average, changes in system parameters are suspected, so μ_1 (which is $< \mu_2$) is chosen and Ω is made larger for good parameter tracking. Else, the error is assumed to be resulting from noise and μ_2 is chosen for smaller Ω , making the algorithm less sensitive to the noise for smooth steady state. ε_{min} is a threshold value used to maintain the values of previous $\hat{\theta}$ and P when the current generated errors are sufficiently small.

The RLAE + VFF algorithm proposed in Eq. (25a) to Eq. (25f) combines the LAE identification that has an IV element to overcome bias problem and the VFF method for fast tracking and smooth steady-state estimation. Next, the property of the proposed LAE + VFF estimation algorithm is investigated.

4. Property of the LAE+VFF algorithm

4.1 Theorem 1 Consider a system described by Eq. (16) with the LAE + VFF estimation and $\delta\theta_k = \theta_k - \theta_{k-1}$, where subscript k represents a discrete-time index. For $\kappa > 0$, $0 < \mu_1 < \mu_2 < 1$ and $\gamma = \mu^T$, we have

$$\|\theta_k - \hat{\theta}_k\| \leq Z_1 \alpha^k + Z_2 \left[\sup_i |\eta_i| + \sup_i |\delta\theta_i| \right] \quad (26)$$

for some constants $Z_1, Z_2 > 0$, $0 < \alpha < 1$ and all k , subject to Assumption 1 and Assumption 2.

Assumption 1

There exist some constants $0 < Z_{lo} < Z_{hi} < \infty$ such that for all k ,

$$0 < Z_{lo} \leq \|\phi_k\|^2 \leq Z_{hi} \quad (27)$$

The upper bound on ϕ_k , in Eq. (27), is guaranteed if the system is stable and noise are bounded. The lower bound only indicates the fact that if ϕ_k is zero or very small, no or very little information on the unknown parameter vector can be extracted.

Assumption 2

The regressor is persistently exciting i.e.

$$\beta_1 I \leq \frac{1}{l_k} \sum_{k=k_0}^{k_0+l_k-1} \phi_k \phi_k^T \leq \beta_2 I \quad (28)$$

for some $0 < \beta_1 \leq \beta_2$, $l_k > 0$ and all $k \geq 0$. This is a standard assumption for the convergence in the adaptive literature.

4.2 Proof of Theorem 1 For the discrete-time LAE + VFF estimator algorithm given in Eq. (25e), let

$$\delta\theta_k = \theta_k - \theta_{k-1}, \text{ indicating parameter change} \quad (29)$$

$$\tilde{\theta}_k = \theta_k - \hat{\theta}_k = \text{estimation error} \quad (30)$$

Multiplying and adding both sides of Eq. (25e) by -1 and θ_k , respectively, we have

$$\underbrace{\theta_k - \hat{\theta}_k}_{\tilde{\theta}_k} = \theta_k - \hat{\theta}_{k-1} - \frac{P_{k-1} \xi_k}{\gamma_k + \phi_k^T P_{k-1} \xi_k} [y_k - \phi_k^T \hat{\theta}_{k-1}] \quad (31)$$

$$\begin{aligned} \tilde{\theta}_k &= \theta_k - \hat{\theta}_{k-1} - \frac{P_{k-1} \xi_k}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \left[\overbrace{\phi_k^T \theta_k + \eta_k}^{y_k} - \phi_k^T \hat{\theta}_{k-1} \right] \\ &= \delta\theta_k + \underbrace{\theta_{k-1} - \hat{\theta}_{k-1}}_{\tilde{\theta}_{k-1}} - \frac{P_{k-1} \xi_k}{\gamma_k + \phi_k^T P_{k-1} \xi_k} [\phi_k^T \theta_k + \eta_k - \phi_k^T \hat{\theta}_{k-1}] \\ &= \delta\theta_k + \tilde{\theta}_{k-1} - \frac{P_{k-1} \xi_k}{\gamma_k + \phi_k^T P_{k-1} \xi_k} [\phi_k^T \theta_k + \eta_k] \end{aligned}$$

$$+ \frac{P_{k-1} \xi_k \phi_k^T (\hat{\theta}_{k-1} - \tilde{\theta}_{k-1})}{\gamma_k + \phi_k^T P_{k-1} \xi_k}$$

$$\begin{aligned} \tilde{\theta}_k &= \delta\theta_k + \tilde{\theta}_{k-1} - \frac{P_{k-1} \xi_k \phi_k^T \theta_k}{\gamma_k + \phi_k^T P_{k-1} \xi_k} - \frac{P_{k-1} \xi_k \eta_k}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \\ &\quad + \frac{P_{k-1} \xi_k \phi_k^T \left(\theta_{k-1} - \delta\theta_k - \tilde{\theta}_{k-1} \right)}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \end{aligned}$$

$$\begin{aligned} \tilde{\theta}_k &= \left(I - \frac{P_{k-1} \xi_k \phi_k^T}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \right) \tilde{\theta}_{k-1} \\ &\quad + \left(I - \frac{P_{k-1} \xi_k \phi_k^T}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \right) \delta\theta_k - \frac{P_{k-1} \xi_k \eta_k}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \end{aligned} \quad (32)$$

Consider the homogeneous part of Eq. (32) i.e.

$$\tilde{\theta}_k = \left(I - \frac{P_{k-1} \xi_k \phi_k^T}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \right) \tilde{\theta}_{k-1} \quad (33)$$

From Equation (22),

$$P_k = \frac{1}{\gamma_k} \left[I - \frac{P_{k-1} \xi_k \phi_k^T}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \right] P_{k-1}, \text{ giving}$$

$$\gamma_k P_k P_{k-1}^{-1} = I - \frac{P_{k-1} \xi_k \phi_k^T}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \quad (34)$$

Notice that the right-hand side term of Equation (34) is similar to the bracketed term of Equation (33). Thus, Equation (33) can be written as

$$\tilde{\theta}_k = \gamma_k P_k P_{k-1}^{-1} \tilde{\theta}_{k-1} \quad (34a)$$

$$(\gamma_k P_k)^{-1} \tilde{\theta}_k = P_{k-1}^{-1} \tilde{\theta}_{k-1} \quad (34b)$$

Introduce a Lyapunov function, V_t ,

$$V_t = \tilde{\theta}_k^T (\gamma_k^{-1} P_k^{-1}) \tilde{\theta}_k \quad (35)$$

Its first difference is,

$$\begin{aligned} V_t - V_{t-1} &= \tilde{\theta}_k^T (\gamma_k^{-1} P_k^{-1}) \tilde{\theta}_k - \tilde{\theta}_{k-1}^T (\gamma_{k-1}^{-1} P_{k-1}^{-1}) \tilde{\theta}_{k-1} \\ &= \tilde{\theta}_k^T P_{k-1}^{-1} \tilde{\theta}_{k-1} - \tilde{\theta}_{k-1}^T (\gamma_{k-1}^{-1} P_{k-1}^{-1}) \tilde{\theta}_{k-1} \\ &= [\tilde{\theta}_k - \gamma_{k-1}^{-1} \tilde{\theta}_{k-1}]^T P_{k-1}^{-1} \tilde{\theta}_{k-1} \\ &= \left[\tilde{\theta}_{k-1} - \underbrace{\frac{P_{k-1} \xi_k \phi_k^T \tilde{\theta}_{k-1}}{\gamma_k + \phi_k^T P_{k-1} \xi_k}}_{\tilde{\theta}_k} - \gamma_{k-1}^{-1} \tilde{\theta}_{k-1} \right]^T P_{k-1}^{-1} \tilde{\theta}_{k-1} \\ &= \tilde{\theta}_{k-1}^T \left[\underbrace{\left(1 - \frac{1}{\gamma_{k-1}} \right) P_{k-1}^{-1}}_{-ve} - \underbrace{\frac{\phi_k \xi_k^T}{\gamma_k + \phi_k^T P_{k-1} \xi_k}}_{+ve} \right] \tilde{\theta}_{k-1} \\ &= \tilde{\theta}_{k-1}^T \underbrace{\left[\left(1 - \frac{1}{\gamma_{k-1}} \right) P_{k-1}^{-1} - \frac{\phi_k \xi_k^T}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \right]}_{Q_{Lyap}} \tilde{\theta}_{k-1} \end{aligned} \quad (36)$$

Since P_{k-1} is a non-singular diagonal matrix with all its main diagonal elements having positive values and $0 < \gamma_k, \gamma_{k-1} < 1$, the term, Q_{Lyap} in Eq. (36) is a negative definite matrix. This concludes that the homogeneous part of Equation (32) is exponentially stable. Also, from Equation (34a),

$$\begin{aligned} \tilde{\theta}_k &= \gamma_k P_k P_{k-1}^{-1} \tilde{\theta}_{k-1} \\ &= \gamma_k P_k P_{k-1}^{-1} (\gamma_{k-1}^{-1} P_{k-1} P_{k-2}^{-1}) \tilde{\theta}_{k-2} \\ &= \gamma_k \gamma_{k-1} P_k P_{k-1}^{-1} P_{k-1} P_{k-2}^{-1} (\gamma_{k-2}^{-1} P_{k-2} P_{k-3}^{-1}) \tilde{\theta}_{k-3} \\ &= \gamma_k \gamma_{k-1} \dots \gamma_0 P_k P_0^{-1} \tilde{\theta}_0 \end{aligned} \quad (37)$$

Using Schwartz inequality and noting that from Assumption 2, P_k^{-1} is uniformly bounded above and below i.e. there exist some constants $0 < \sigma_1 < \sigma_2$ so that $\sigma_1 I \leq P_k^{-1} \leq \sigma_2 I$, from Eq. (37),

$$\begin{aligned} \|\tilde{\theta}_k\| &\leq |\gamma_k| |\gamma_{k-1}| \dots |\gamma_0| \|P_k\| \|P_0^{-1}\| \|\tilde{\theta}_0\| \\ &\leq \gamma^{-k} \|P_k\| \|P_0^{-1}\| \|\tilde{\theta}_0\| \\ &\leq Z \gamma^{-k} \|\tilde{\theta}_0\| \\ &\leq Z_1 \alpha^k + Z_2 \left[\sup_i |\eta_i| + \sup_i |\delta \theta_i| \right] \end{aligned}$$

5. Simulation

To evaluate the effectiveness of the LAE+VFF method, the algorithm is applied to a solid-axle railway wheelset system. The wheelset consists of two coned/profiled wheels rigidly connected by an axle. Two of the wheelset parameters can vary significantly: the ‘creep coefficient’, f , and the ‘conicity’, λ . f is a parameter that is nonlinearly dependent on the normal force between wheel and rail, and its value can vary between 5 – 10 MN, whereas λ , which is a term related to the coning of the wheel tread, in practice has its value

varying between 0.05 and 0.5.

The solid-axle wheelset suspended to ground via lateral springs and dampers as shown in Fig. 2 is considered.

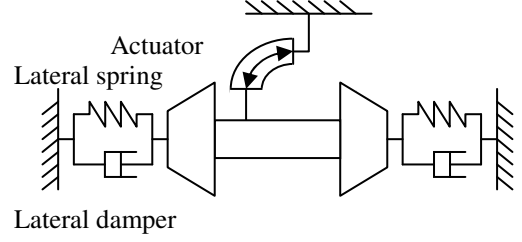


Fig. 2: Solid-axle wheelset arrangement for simulation

The wheelset is assumed to be traveling at 83 m/s along a straight track with lateral irregularities. It is controlled using an active yaw damping technique, in which a rotary actuator provides the yaw torque. The control input signal is provided by a linear quadratic regulator (LQR) designed to minimise the wheelset’s lateral displacement relative to the track. The transfer function of the wheelset system, whose output is taken as the wheelset’s lateral velocity, is given by Eq. (38). Symbols and values used are given in the Appendix.

$$G(s) = \frac{b_2 s^2}{s^5 + a_4 s^4 + a_3 s^3 + a_2 s^2 + a_1 s + a_0} \quad (38)$$

where

$$\begin{aligned} a_4 &= \frac{2fl^2}{Iv} + \frac{2f}{mv} + \frac{C_l}{m}, \quad a_3 = \frac{4f^2 l^2}{mIv^2} + \frac{2fl^2 C_l}{mlv} + \frac{K_l}{m} \\ a_2 &= \frac{2fl^2 K_l}{mlv}, \quad a_1 = \frac{4f^2 l \lambda}{mlr_o}, \quad a_0 = 0, \quad b_2 = \frac{2f}{ml} \end{aligned}$$

The LAE+VFF estimator was used to estimate a_i and b_i . Then $\hat{\lambda}$ and \hat{f} were calculated from any suitable combination of b_2 to a_4 . Monte Carlo simulations have been done to investigate the effectiveness of the method. Firstly, the normal recursive least-squares error (RLSE) [2] and LAE estimation methods using fixed forgetting factor i.e. $\mu_1 = \mu_2 = 0.99$, are compared. $f_s = 1/T_s = 300$ Hz, $l_f = 3$, $\kappa = e_{min} = 0.01$ are used. The RLSE algorithm is similar to LAE algorithm but with $\xi_k = \phi_k$. Fig. 3 shows the parameter tracking performance of both estimators. It can be seen that the LAE estimator tracks the changes in system parameters much more quickly. Moreover, the estimate of λ using RLSE contains significant bias whereas the estimates of both f and λ using LAE estimator converge to true values.

For the LAE+VFF, $\mu_1 = 0.80$, $\mu_2 = 0.99$ have been used. Fig. 4 shows that LAE+VFF estimator takes less than a second to reach the new value of f compared to the LAE estimator that takes longer time and Table 1 shows that LAE+VFF estimator is superior to both RLSE and LAE.

Table 2 shows the effect of the choices of f_s and l_f values on the effectiveness of the LAE+VFF method. It is observed from Table 2 that proper choices of f_s and l_f are important for good estimation results, and the pair

$f_s=300$ Hz and $l_f=3$ gives the best estimates because with these values, the frequency response of the LIF then lies exactly between the least stable (with $f=5$ MN, $\lambda=0.05$) and the most stable (with $f=10$ MN, $\lambda=0.5$) wheelset system's frequency bands, as shown in Fig. 5.

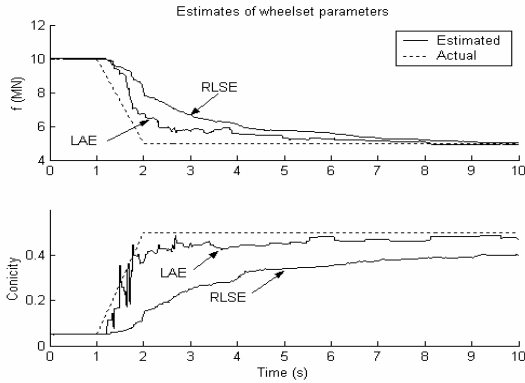


Fig. 3: RLSE versus LAE using fixed forgetting factor

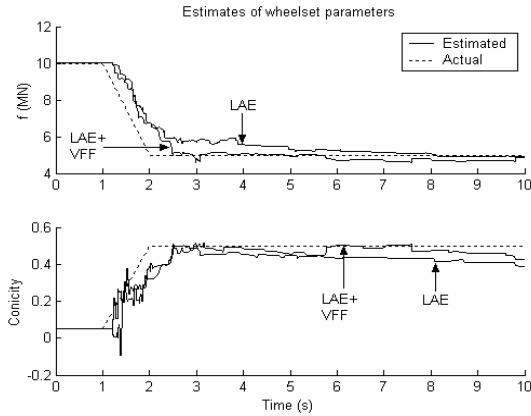


Fig. 4: LAE versus LAE+VFF

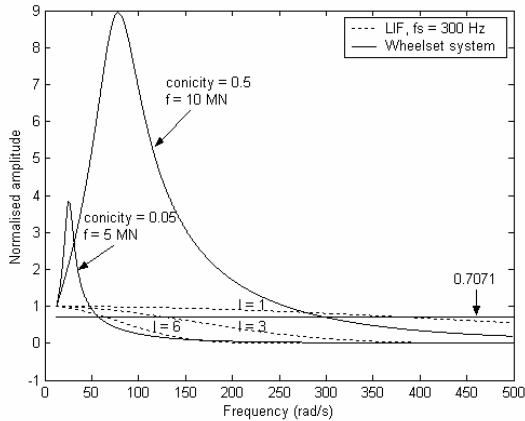


Fig. 5: Frequency responses of the LIF and the wheelset system

Table 1: Average estimation error (A.E.E) of RLSE, LAE & LAE+VFF estimators

	RLSE	LAE	LAE+VFF
\hat{f}	13.7±1.2 %	7.8±0.8 %	5.58±1.0 %
$\hat{\lambda}$	32.5±2.5 %	15.7±1.3 %	9.72±1.4 %

Table 2: Effects of different combinations of f_s and l_f on the estimates for LAE+VFF

l_f	$f_s = 300$ Hz		$l_f = 3$		
	A.E.E (%)		f_s , Hz	A.E.E (%)	
	\hat{f}	$\hat{\lambda}$		\hat{f}	$\hat{\lambda}$
1	17.5±2.5	30.0±2.1	100	28.9±2.8	43.1±3.2
3	5.5±1.0	9.7±1.4	300	5.5±1.0	9.7±1.4
8	36.3±3.3	48.8±4.0	800	23.4±3.2	21.4±1.2

6. Conclusions

The LAE with fixed forgetting factor gives better parameter estimates compared to the RLSE, whereas the LAE+VFF offers even better estimation and tracking of system parameters that are subject to abrupt changes, provided that the f_s and l_f values are chosen accordingly. It has also been proven that the estimation error of the proposed LAE+VFF estimation algorithm is bounded.

Acknowledgement

The authors would like to thank Universiti Teknologi Malaysia for their support.

References

- [1] Gawthrop, P.J., "A Continuous-Time Approach To Discrete-Time Self-Tuning Control", Opt. Control Appl. Meth., Vol. 3, (1982), pp. 399–414
- [2] Stjerc, V., "Least Squares Parameter Estimation", Automatica, Vol. 16, (1980), pp. 535–550
- [3] Van den Hof, P.M.J. and Schrama, J.P., "An Indirect Method For Transfer Function Estimation From Closed-Loop Data", Automatica, Vol. 29, No. 6, (1993), pp. 1523–1527
- [4] Young, P.C., "Instrumental Variable Method For Real-Time Identification Of A Noisy Process", Automatica, Vol. 6, (1970), pp. 271–287
- [5] Unbehauen, H. and Rao, G.P., "Continuous-Time Approaches To System Identification – A Survey. Automatica, Vol. 26, No. 1, (1990), pp. 23–35
- [6] Sagara, S. and Zhao, Z-Y, "Numerical Integration Approach To On-Line Identification Of Continuous-Time Systems," Automatica, Vol. 26, No. 1, (1990), pp. 63–74
- [7] Kowalczyk, Z., and Kozłowski, J., "Integrated Squared Error And Integrated Absolute Error In Recursive Identification Of Continuous-Time Plants", Proc. UKACC Int. Conf. On Control, Vol. 455, (1998), pp. 693 – 698
- [8] Bai, E-W. and Huang, Y-F, "Variable Gain Parameter Estimation Algorithms For Fast Tracking and Smooth Steady State", Automatica, Vol. 36, (2000), pp. 1001 – 1008

Appendix

Symbol	Parameter	Value
l	Half-gauge length	0.7 m
r_o	Wheel radius	0.45 m
C_b	Lateral damping per wheelset	50 kN/s/m
K_l	Lateral stiffness per wheelset	200 kN/m
I	Wheelset inertia	700 kg m ²
m	Wheelset mass	1250 kg
v	Forward speed	83.3 m/s