RAILWAY WHEELSET PARAMETER ESTIMATION USING SIGNALS FROM LATERAL VELOCITY SENSOR

H. Selamat§, A. J. Alimin‡ & Y. M. Sam§
§Electrical Engineering Faculty
Universiti Teknologi Malaysia
81310 Skudai, Johor, Malaysia
‡Universiti Tun Hussein Onn Malaysia
86400 Parit Raja, Batu Pahat, Johor, Malaysia
Emails: hazlina@fke.utm.my, ajais@uthm.edu.my, yahaya@fke.utm.my

Abstract- A type of parameter estimation technique based on the linear integral filter (LIF) method, the least-absolute error with variable forgetting factor (LAE+VFF) estimation method, is proposed in this paper to estimate the railway wheelset parameters modelled as a time-varying continuous-time (C-T) system. The inputs to the parameter estimator are the control signal and the railway wheelset system output, which is the wheelset’s lateral velocity. 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. Simulation results have shown that the LAE with fixed forgetting factor gives better parameter estimates compared to the recursive least-squares error (RLSE) method, whereas the LAE+VFF offers even better estimation and tracking of system parameters that are subject to abrupt changes, provided that the $f_i$ 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.

Index terms: linear integral filter; least-absolute error; variable forgetting factor; continuous-time estimation; solid-axle wheelset
I. 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.

II. SYSTEM DESCRIPTION

A C-T model described by Equation 1 is considered.

\[
G(s) = \frac{b_ms^n + b_{m-1}s^{n-1} + \ldots + b_1s + b_0}{s^n + a_{n-1}s^{n-1} + \ldots + a_1s + a_0} = \frac{Y(s)}{U(s)}
\]  (1)

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

\[
y_n(t) = \begin{bmatrix} y_{n-1}(t) & \ldots & -y(t) & u_m(t) & \ldots & u(t) \end{bmatrix} \begin{bmatrix} a_{n-1} & \ldots & a_0 & b_m & \ldots & b_0 \end{bmatrix}^T
\]  (2)

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

Including \( \eta(t) \), which represents the measurement noise affecting the output, Equation 2 can then be written in the following regression model form:
\[ y_{\text{diff}}(t) = \phi_{\text{diff}}^T(t)\theta + \eta(t) \]  

where

\[ y_{\text{diff}}(t) = y_n(t) \]

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

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

To estimate \( \theta \) 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 (Equation 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 (\( \mathcal{L}\mathcal{D} \)) and the operator’s outputs can be used directly in the estimation algorithm as for D-T systems (Figure 1).

![Figure 1: Linear dynamic operation for C-T parameter estimation](image)

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 \( n^{\text{th}} \)-order integration operator, \( \Gamma^n \), is introduced and
multiple integrations over the time interval \([t-l_f T, t]\) are performed on both sides of Equation 3, giving Equation 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).

\[
y(t) = \phi^T(t) \theta + \eta(t)
\]  

where,

\[
y(t) = \Gamma^n y_n(t)
\]

\[
\phi^T(t) = \begin{bmatrix} -\Gamma^n y_{n-1}(t) & \cdots & -\Gamma^n y(t) & \Gamma^n u_m(t) & \cdots & \Gamma^n u(t) \end{bmatrix}
\]

\[
\theta(t) = \begin{bmatrix} a_{n-1} & \cdots & a_0 & b_m & \cdots & b_0 \end{bmatrix}^T
\]

The C-T multiple integrals in the regression vector, \(\phi(t)\), of Equation 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 Equation 4.

III. LEAST-ABSOLUTE ERROR WITH VARIABLE FORGETTING FACTOR (LAE+VFF) ESTIMATION

a. Continuous-time LAE+VFF

The C-T LAE + VFF estimation method minimises the criterion given by Equation 5.

\[
J_{LAE}(\theta) = \int_0^t \mu^{t-\tau} |\varepsilon(\tau)| d\tau
\]

where \(\varepsilon(\tau) = y(\tau) - \phi^T(\tau) \theta\), \(t\) and \(\tau\) are normalised quantities of the same independent time variable and \(\mu\) is the C-T VFF that has a value between 0 and 1. Minimisation of Equation 5 with respect to \(\theta\) gives

\[
\frac{\partial J_{LAE}(\theta)}{\partial \theta} = \int_0^t \mu^{t-\tau} \phi^T(\tau) \frac{\varepsilon(\tau)}{|\varepsilon(\tau)|} d\tau = 0
\]

where \(\varepsilon(\tau) = y(\tau) - \phi^T(\tau) \theta\). Since \(\varepsilon(t)\) is not known \textit{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 Equation 6 can be derived as
\( \hat{\theta}(t) = M^{-1}(t)N(t) \) \hspace{1cm} (7)

\[
M(t) = \int_{0}^{t} \mu^{-\tau} \frac{\phi_{\tau}(\tau)\phi_{\tau}^{T}(\tau)}{[\tilde{e}(\tau)]} \, d\tau
\] \hspace{1cm} (8)

\[
N(t) = \int_{0}^{t} \mu^{-\tau} \frac{\phi_{\tau}(\tau)y_{\tau}(\tau)}{[\tilde{e}(\tau)]} \, d\tau
\] \hspace{1cm} (9)

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

\[
\tilde{P}(t) = \frac{d}{dt} M^{-1}(t) = -\tilde{P}(t) \ln \mu - \frac{\tilde{P}(t)\phi_{\tau}(t)\phi_{\tau}^{T}(t)\tilde{P}(t)}{[\tilde{e}(t)]}
\] \hspace{1cm} (10)

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

\[
\dot{\hat{\theta}}(t) = \frac{d}{dt} [\tilde{P}(t)N(t)] = \frac{\tilde{P}(t)\phi_{\tau}(t)e(t)}{[\tilde{e}(t)]}
\] \hspace{1cm} (11)

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

(i) Model error is evaluated

\[
e(t) = y(t) - \phi^{T}(t)\hat{\theta}(t)
\] \hspace{1cm} (12)

(ii) Parameter covariance matrix is updated (from Equation 10)

\[
d\tilde{P}(t) = -\left[\tilde{P}(t) \ln \mu + \tilde{P}(t)\phi(t)\frac{\phi(t)\phi_{\tau}^{T}(t)\tilde{P}(t)}{[\tilde{e}(t)]}\right]dt
\] \hspace{1cm} (13)

(iii) The estimates are corrected (from Equation 11)

\[
d\hat{\theta}(t) = \tilde{P}(t)\frac{\phi(t)}{[\tilde{e}(t)]}e(t)\,dt
\] \hspace{1cm} (14)

b. Discretisation of the regression vector

An \( n^{th} \)-order multiple integration, \( \Gamma^{n} \), of a C-T signal, \( x(t) \), over the time interval \([t-1T, t)\) yields

\[
\Gamma^{n}x(t) = \left(1 - e^{-s/T}\right)^{n}x(t) = \int_{t-1T}^{t} \cdots \int_{t_{n-1}T}^{t_{n}T} x(t) \, dt_{n} \cdots dt_{1}
\]

Thus, the elements of \( \phi_{\tau}(t) \) can be evaluated using an operator polynomial, \( J^{n}_{i}(q^{-1}) \) \[7\] as Equation 15.
\[ \Gamma^n x_i(t) = \left(1 - e^{-s/T}\right)^n \frac{d^n x(t)}{dt^n} = J^n_i(q^{-1})x(kT) \] (15)

where

\[ J^n_i(q^{-1}) = \frac{a!T^{n-i}}{(a+n-i)!} \left[D_i(q^{-1})\right]^{b-i} : N_{a+n-i}(q^{-1}) \times N_0(q^{-1}) \cdot \left(1 - q^{-l_f}\right) \]

\[ D_i(q^{-1}) = 1 + q^{-1} + \ldots + q^{-(l_f-1)} = \frac{1 - q^{-l_f}}{1 - q^{-1}} \]

\[ q^{-1} = \text{unit - delay operator} \]

\[ l_f = \text{length of the discrete - time integration window} \]

\[ i = 0, 1, \ldots, n \]

\[ n = \text{system order} \]

\[ a, b = \text{orders of the input and output signal interpolation} \]

\[ N_0(q^{-1}) = N_1(q^{-1}) = 1 \]

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

\[ w_{b,1} = w_{b,b} = 1 \]

\[ w_{b,j} = (b+1-j)w_{b-1,j-1} + jw_{b-1,j} \]

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

\[ y(k) = \phi^T(k)\theta + \eta(k) \] (16)

where,

\[ y(k) = J^n (q^{-1})y(k) \]

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

\( T \) is generally selected according to \( \frac{\pi}{150\omega_n} < T < \frac{\pi}{10\omega_n} \), where \( \omega_n \) 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.
c. Discretisation of the estimator

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

\[
M(k) \approx T \sum_{i=1}^{k} \mu^{(k_i)} \bar{\xi}(i) \phi^T(i) = \gamma M(k-1) + T \bar{\xi}(k) \phi^T(k) \tag{17}
\]

\[
N(k) \approx T \sum_{i=1}^{k} \mu^{(k_i)} \bar{\xi}(i) y(i) = \gamma N(k-1) + T \bar{\xi}(k) y(k) \tag{18}
\]

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

where \( \bar{\xi}(k) = \frac{\phi(k)}{\bar{\xi}(k)} \) and \( \gamma = \mu^T \) i.e the \( T \)-power of the continuous-time VFF, \( \mu \), representing its discrete-time equivalent. Applying Matrix Inversion Lemma (MIL) to Equation 17, \( M^{-1}(t) = \bar{P}(t) \) can be evaluated as Equation 20.

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

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

\[
\dot{\theta}(k) = \hat{\theta}(k-1) + P(k) \bar{\xi}(k) \varepsilon(k)
\]

\[
= \hat{\theta}(k-1) + \frac{P(k-1)}{\gamma + \phi(k) P(k-1) \bar{\xi}(k)} \bar{\xi}(k) \varepsilon(k) \tag{21}
\]

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

Equation 21 and Equation 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

\[
\dot{\theta}(k) = \hat{\theta}(k-1) + \Omega(k) \bar{\xi}(k) \varepsilon(k) \tag{23}
\]

where

\[
\varepsilon(k) = y(k) - \phi^T(k) \hat{\theta}(k-1)
\]
\[ \xi(t) = \frac{\phi(t)}{|\xi(t)|} = \left( y(k) - \phi^T(k)\hat{\theta}(k) \right) = \text{IV vector} \quad \Omega(k) = \frac{P(k)}{\gamma + \phi^T(k)P(k)\xi(k)} = \text{gain matrix} \]

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

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

To track time-varying parameters, the choice of \( \Omega \) in Equation 23 is crucial. Large \( \Omega \) results in good parameter tracking but the estimator becomes more sensitive to noise. In contrast, small \( \Omega \) 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 \textit{a priori} prediction error

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

(ii) Calculate the \textit{a posteriori} prediction error

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

(iii) Evaluate the IV vector,

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

(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 \( \gamma \), which is the D-T equivalent of the C-T VFF, \( \mu \)

\[ \mu = \begin{cases} \mu_1 & \text{if } |\varepsilon(k)| > \kappa \\ \mu_2 & \text{if } |\varepsilon(k)| \leq \kappa \end{cases} \]

\[ \gamma(k) = \mu^T \quad (25f) \]
\( \kappa \) 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 \( \kappa \) 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 \( \mu_1 \) (which is \( < \mu_2 \)) is chosen and \( \Omega \) is made larger for good parameter tracking. Else, the error is assumed to be resulting from noise and \( \mu_2 \) is chosen for smaller \( \Omega \), making the algorithm less sensitive to the noise for smooth steady state. \( \varepsilon_{\text{min}} \) 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 Equation 25(a) to Equation 25(f) 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.

IV. PROPERTY OF THE LAE + VFF ALGORITHM

a. Theorem 1

Consider a system described by Equation 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

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

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 \( k \),

\[
0 < Z_{lo} \leq \| \phi_k \|^2 \leq Z_{hi}
\]

The upper bound on \( \phi_k \), in Equation 27, is guaranteed if the system is stable and noise are bounded. The lower bound only indicates the fact that if \( \phi_k \) 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.

$$\beta IL \leq \frac{1}{I} \sum_{k=1}^{k-1} \phi_k \phi_k^T \leq \beta_2 I$$

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

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

$$\delta \theta_k = \theta_k - \theta_{k-1},$$

indicating parameter change

$$\tilde{\theta}_k = \theta_k - \tilde{\theta}_k$$

estimation error

Multiplying and adding both sides of Equation 25(e) by $-1$ and $\theta_k$, respectively, we have

$$\frac{\theta_k - \tilde{\theta}_k}{\tilde{\theta}_k} = \theta_k - \tilde{\theta}_k - \frac{P_{k-1} \xi_k}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \left[ y_k - \phi^T \tilde{\theta}_{k-1} \right]$$

$$\tilde{\theta}_k = \theta_k - \tilde{\theta}_k - \frac{P_{k-1} \xi_k}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \left[ \phi_k^T \theta_k + \eta_k - \phi_k^T \tilde{\theta}_{k-1} \right]$$

$$= \delta \theta_k + \theta_k - \tilde{\theta}_k - \frac{P_{k-1} \xi_k}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \left[ \phi_k^T \theta_k + \eta_k - \phi_k^T \tilde{\theta}_{k-1} \right]$$

$$= \delta \theta_k + \tilde{\theta}_k - \frac{P_{k-1} \xi_k}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \left[ \phi_k^T \theta_k + \eta_k \right] + \frac{P_{k-1} \xi_k}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \phi_k^T \left( \theta_k - \tilde{\theta}_k - \tilde{\theta}_{k-1} \right)$$

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

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

Consider the homogeneous part of Equation 32 i.e.

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

(33)
From Equation 22,

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

giving

\[ \gamma_k P_k P_{k-1}^{-1} = I - \frac{P_{k-1} \xi_k \phi_k^T}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \]

(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

\[ \vec{\theta}_k = \gamma_k P_k P_{k-1}^{-1} \vec{\theta}_{k-1} \]  
(34a)

\[ (\gamma_k P_k)^{-1} \vec{\theta}_k = P_{k-1}^{-1} \vec{\theta}_{k-1} \]  
(34b)

Introduce a Lyapunov function, \( V \),

\[ V = \vec{\theta}_k^T (\gamma_k^{-1} P_k^{-1}) \vec{\theta}_k \]  
(35)

Its first difference is,

\[ V_t - V_{t-1} = \vec{\theta}_k^T (\gamma_k^{-1} P_k^{-1}) \vec{\theta}_k - \vec{\theta}_{k-1}^T (\gamma_{k-1}^{-1} P_{k-1}^{-1}) \vec{\theta}_{k-1} \]

\[ = \vec{\theta}_k^T P_{k-1}^{-1} \vec{\theta}_{k-1} - \vec{\theta}_{k-1}^T (\gamma_{k-1}^{-1} P_{k-1}^{-1}) \vec{\theta}_{k-1} \]

\[ = [\vec{\theta}_k - \gamma_{k-1}^{-1} \vec{\theta}_{k-1}]^T P_{k-1}^{-1} \vec{\theta}_{k-1} \]

\[ = \left[ \vec{\theta}_{k-1} - \frac{P_{k-1} \xi_k \phi_k^T \vec{\theta}_{k-1}}{\gamma_k + \phi_k^T P_{k-1} \xi_k} - \gamma_{k-1}^{-1} \vec{\theta}_{k-1} \right]^T P_{k-1}^{-1} \vec{\theta}_{k-1} \]

\[ = \vec{\theta}_{k-1}^T \left[ 1 - \frac{1}{\gamma_{k-1}} \right] P_{k-1}^{-1} - \frac{\phi_k^T \xi_k}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \vec{\theta}_{k-1} \]

\[ = \vec{\theta}_{k-1}^T \left[ \frac{1 - \frac{1}{\gamma_{k-1}}}{\gamma_{k-1}} \right] P_{k-1}^{-1} - \frac{\phi_k^T \xi_k}{\gamma_k + \phi_k^T P_{k-1} \xi_k} \vec{\theta}_{k-1} \]

(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 \( Q_{\text{Lyap}} \) in Equation 36 is a negative definite matrix. This concludes that the homogeneous part of Equation 32 is exponentially stable. Also, from Equation 34(a),
\[ \tilde{\theta}_k = \gamma_k P_k P_{k-1}^{-1} \tilde{\theta}_{k-1} \]
\[ = \gamma_k P_k P_{k-1}^{-1} \left( \gamma_{k-1} P_{k-1} P_{k-2}^{-1} \right) \tilde{\theta}_{k-2} \]
\[ = \gamma_k \gamma_{k-1} P_k P_{k-1}^{-1} P_{k-1} P_{k-2}^{-1} \left( \gamma_{k-2} P_{k-2} P_{k-3}^{-1} \right) \tilde{\theta}_{k-3} \]  

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 Equation 37,

\[
\left\| \tilde{\theta}_k \right\| \leq |\gamma_k| |\gamma_{k-1}| \cdots |\gamma_0| \left\| P_k P_{k-1}^{-1} \tilde{\theta}_0 \right\| \\
\leq \gamma^{-k} \left\| P_k P_{k-1}^{-1} \tilde{\theta}_0 \right\| \\
\leq Z \gamma^{-k} \left\| \tilde{\theta}_0 \right\| \\
\leq Z_1 |\alpha|^k + Z_2 \left( \sup_i |\eta_i| + \sup_i |\delta \theta_i| \right) 
\]

V. 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’, \( \lambda \). \( f \) 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 \( \lambda \), 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 Figure 2 is considered.

![Figure 2: Solid-axle wheelset arrangement for simulation](image)
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 Equation 38. Symbols and values used are given in the Appendix.

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

(38)

where

\[ a_4 = \frac{2 fl^2}{Iv} + \frac{2 f}{mv} + \frac{C_i}{m}, a_3 = \frac{4 f^2 l^2}{mlv^2} + \frac{2 fl^2 C_i}{mlv} + \frac{K_J}{m}, a_2 = \frac{2 f^2 K_I}{mlv}, a_1 = \frac{4 f^2 l \lambda}{mlr_o}, \]

\[ a_0 = 0 \quad \text{and} \quad b_2 = \frac{2 f}{ml}. \]

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. \( \mu_1 = \mu_2 = 0.99 \), were compared. \( f_s = 1/T_s = 300 \) Hz, \( l_f = 3 \), \( \kappa = e_{\min} = 0.01 \) were used. The RLSE algorithm is similar to LAE algorithm but with \( \xi_k = \phi_k \). Figure 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 \( \lambda \) using RLSE contains significant bias whereas the estimates of both \( f \) and \( \lambda \) using LAE estimator converge to true values. For the LAE+VFF, \( \mu_1 = 0.80, \mu_2 = 0.99 \) have been used. Figure 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 Figure 5.
Figure 3: RLSE versus LAE using fixed forgetting factor

Figure 4: LAE versus LAE+VFF
Table 1: Average estimation error (A.E.E) of RLSE, LAE & LAE+VFF estimators

<table>
<thead>
<tr>
<th></th>
<th>RLSE</th>
<th>LAE</th>
<th>LAE+VFF</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{f} )</td>
<td>13.7±1.2 %</td>
<td>7.8±0.8 %</td>
<td>5.58±1.0 %</td>
</tr>
<tr>
<td>( \hat{\lambda} )</td>
<td>32.5±2.5 %</td>
<td>15.7±1.3 %</td>
<td>9.72±1.4 %</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>( l_f )</th>
<th>( f_s ) (Hz)</th>
<th>A.E.E (%)</th>
<th>( l_f )</th>
<th>( f_s ) (Hz)</th>
<th>A.E.E (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>17.5±2.5</td>
<td>5.5±1.0</td>
<td>300</td>
<td>5.5±1.0</td>
</tr>
<tr>
<td>3</td>
<td>5.5±1.0</td>
<td>30.0±2.1</td>
<td>9.7±1.4</td>
<td>300</td>
<td>9.7±1.4</td>
</tr>
<tr>
<td>8</td>
<td>800</td>
<td>36.3±3.3</td>
<td>48.8±4.0</td>
<td>800</td>
<td>36.3±3.3</td>
</tr>
</tbody>
</table>

Figure 4: LAE versus LAE+VFF
VI. 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 $f_i$ 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 and Universiti Tun Hussein Onn Malaysia for their supports.

APPENDIX

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$l$</td>
<td>Half-gauge length</td>
<td>0.7 m</td>
</tr>
<tr>
<td>$r_o$</td>
<td>Wheel radius</td>
<td>0.45 m</td>
</tr>
<tr>
<td>$C_b$</td>
<td>Lateral damping per wheelset</td>
<td>50 kNs/m</td>
</tr>
<tr>
<td>$K_l$</td>
<td>Lateral stiffness per wheelset</td>
<td>200 kN/m</td>
</tr>
<tr>
<td>$I$</td>
<td>Wheelset inertia</td>
<td>700 kg m²</td>
</tr>
<tr>
<td>$m$</td>
<td>Wheelset mass</td>
<td>1250 kg</td>
</tr>
<tr>
<td>$v$</td>
<td>Forward speed</td>
<td>83.3 m/s</td>
</tr>
</tbody>
</table>

REFERENCES


