Fast Model-Based Extremum Seeking on Hammerstein Plants

Jalil Sharafi, William H. Moase, Rohan C. Shekhar and Chris Manzie

Abstract—Partial plant knowledge may be used to develop model-based extremum seeking algorithms, however existing results rely on a type of time-scale separation which leads to slow optimization relative to the plant dynamics. In this work, a fast model-based extremum seeking scheme is proposed for a Hammerstein plant, and semi-global stability results are provided. Simulation results are used to validate the theoretical results.

I. INTRODUCTION

In many control applications, it is desirable to optimize the steady-state output of a plant, despite the plant being partially or completely unknown. Techniques used to achieve this goal are called Extremum Seeking (ES).

In a widely used class of ES, the plant dynamics and its steady state behavior are assumed to be completely unknown (black-box). The steady-state input-output map, \( y = f(u) \), is typically assumed to have a unique minimum at \( u^* \). Krstić and Wang provided the first rigorous stability proof for a class of ES methods that deal with these problems [1]. These methods employ a low frequency dither signal to estimate the local gradient, \( f'(u) \), which is then used in a gradient descent fashion to optimize the output. The local stability result in [1] appears as a precursor for a series of theoretical advances, including: a semi-global stability result [3], multi-input ES [2], stochastic perturbation-based ES [4] and discrete-time ES [5]. All of these methods can be well described in a unified manner as an interconnection of a function derivative estimator in series with an optimization algorithm—Partially there is a trade-off between the size of the dither frequency and the rate of convergence. There are a few results in the literature which allows any nonlinearity within the plant to be represented within a static mapping which is “sandwiched” between Linear Time Invariant (LTI) input and output dynamics. Krstić [7] gave a local stability result for WH plants with an unknown quadratic nonlinearity and known input and output dynamics. Moase and Manzie [8] later showed that it was possible to achieve semi-global stability for WH plants with fairly arbitrary static nonlinearities and requiring little more knowledge than the relative degree of the plant’s input and output dynamics. This is achieved through appropriate design of filters applied to the plant’s input and output.

Unlike traditional black-box approaches, in some applications it is possible to get some knowledge about the plant steady-state behavior, \( f(\cdot) \), either based on previous open-loop tests or physical modeling [9], [10]. In these situations, the plant steady-state map is a known function of the input, but is parameterized by some unknown parameters \( \theta \), in the form \( f(u, \theta) \). Hence, instead of estimating the gradient, an unknown parameter vector is estimated and used with a model-based optimization law to achieve ES. These methods are often referred to as model-based ES [11] or grey-box ES [12]. Recently Nesić et al. have proposed a systematic way to design model-based ES for fairly general nonlinear plants [12]. This method requires time-scale separation of the plant dynamics, parameter estimator and optimizer in a similar manner to traditional black-box approaches, and consequently this tuning strategy generally exhibits slow convergence.

This paper addresses a gap in the literature by proposing a fast model-based ES algorithm. The motivation for this work is to benefit from using a high dither frequency in a model-based ES design. To that end, a single-input single-output (SISO) Hammerstein plant model is considered. A Hammerstein plant model is established as a good model for many engineering systems, particularly for those where the dynamics are attributed to a slow sensor response [13], [14]. The proposed ES scheme uses a gradient algorithm for the parameter estimation and gradient descent for optimization. Semi-global stability is proven, and it is shown that the proposed ES scheme can achieve fast convergence rates by employing a sufficiently large dither frequency. The proposed scheme, however, relies upon strong assumptions on the nonlinearity (generic to model-based approaches) and the plant dynamics structure. Also the persistence of excitation criterion, arising from the parametric nature of the approach, adds more complexity to analysis and design of the ES.

1) Mathematical Preliminaries: Before introducing the main concept, it is convenient to review some preliminaries that will be used throughout the analysis. The set of real numbers is denoted by \( \mathbb{R} \), and \( u(t) \) is a measurable function taking values in \( \mathbb{R} \) for all \( t \geq 0 \).
numbers is denoted by \( \mathbb{R} \). The continuous function \( \alpha : [0, a) \rightarrow \mathbb{R}_{\geq 0} \) is said to be class \( K \) if it is nondecreasing and \( \alpha(0) = 0 \). The function \( \beta : [0, a) \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0} \) is said to be class \( KL \) function if it is of class \( K \) in its first argument and strictly decreasing to zero in its second argument. Denote the \( L_2 \) norm by \( \| \cdot \| \). A ball in \( \mathbb{R}^n \) is defined as \( B^n(r) = \{ x \in \mathbb{R}^n : \| x \| \leq r \} \). Finally, a quantity, \( x \in \mathbb{R}^n \), is \( O(c) \) if there exist \( (c^*, K) \in \mathbb{R}_{\geq 0}^2 \) such that \( \| x \| \leq Kc \) for all \( c \in (0, c^*) \).

II. SYSTEM DESCRIPTION

Fig. 1 provides a schematic diagram of the proposed model-based ES controller. The ES scheme consists of a parameter estimator followed by an optimization law. In this scheme, \( \hat{\theta} \) is the current estimate of the unknown parameter vector. The dither signal, \( a \sin(\Omega t) \), is superimposed on the optimizer output to provide persistency of excitation in the unknown parameters: \( u = \bar{u} + a \sin(\Omega t) \).

A. Plant

Assumption 1. The plant can be represented as a time-invariant SISO mapping, followed by a bi-proper stable linear time invariant system \( F_0(s) \). The static map is considered to be a known function of the input, linearly dependent on some unknown parameters. It can be written in the form

\[
z = f(u)^T \theta, \tag{1}
\]

where \( f(\cdot) : \mathbb{R} \rightarrow \mathbb{R}^{n_u} \) is a known vector of input functions and \( \theta \in \mathbb{R}^{n_u} \) is a vector of unknown parameters. The output filter, \( F_0(s) \), is unknown and has a minimal state space representation

\[
\dot{x} = Ax + Bz, \tag{2a}
\]

\[
y = Cx + Dz, \tag{2b}
\]

where \( x \in \mathbb{R}^{n_x} \), \( z \in \mathbb{R} \), and \( D \in \mathbb{R}_{>0} \).

Remark 1. Equation (1) differentiates between the black-box and the model based ES approaches. Also considering a bi-proper \( F_0(s) \) is made to simplify the presentation of stability analysis in this paper.

Assumption 2. For any given (but unknown) \( \theta \in \Omega_\theta \subset \mathbb{R}^{n_\theta} \) the static map (1) has a unique minimum at \( u^*(\theta) \).

Remark 2. The goal of the proposed ES is to regulate the input as close to \( u^* \) as possible.

Assumption 3. There exist \((v, b) \in \mathbb{R}_{\geq 0}^2 \) such that for any \( \hat{u} \in B^3(v) \) the following holds:

- \( f(\hat{u} + u^*) \) is continuously differentiable sufficiently many times;
- \( 0 < \hat{u} f'(\hat{u} + u^*)^T \theta < b \hat{u}^2 \) (\( \hat{u} \neq 0 \)).

B. Optimizer

The optimizer uses the following gradient descent algorithm

\[
\dot{\hat{u}} = -k\Omega f'(\hat{u})^T \hat{\theta}, \tag{3}
\]

where \( k\Omega \) is the optimizer gain. By selecting sufficiently small \( k \), the output of the optimizer, \( \hat{u} \), is approximately stationary with respect to the signal, \( a \sin(\Omega t) \), let \( \overline{f(\cdot)} \) represent the mean value of \( f(\cdot) \) over one period \( 2\pi/\Omega \):

\[
\overline{f(\hat{u})} := \frac{\Omega}{2\pi} \int_0^{2\pi/\Omega} f(\hat{u} + a \sin(\Omega t)) dt. \tag{4}
\]

C. Estimator

To generate a measure of parameter estimation error, it is necessary to calculate an estimate of the output, \( \hat{y} \), based on the current estimated parameter vector, \( \hat{\theta} \). The predicted output is represented as

\[
\hat{y} = \hat{y}_0 + [f(u) - \overline{f(\hat{u})}]^T \hat{\theta}. \tag{5}
\]

In this estimation model, \( \hat{y}_0 \) represents some “mean” value of the plant output, \( y \), and the remainder represents some fluctuating part of \( y \). In this paper, the gradient descent algorithm is also used for the estimator. Defining the estimation error as \( e := y - \hat{y} \) leads to:

\[
\begin{bmatrix}
\hat{y}_0 \\
\hat{\theta}
\end{bmatrix} = -\Omega \begin{bmatrix} 0 & \partial e^2/2 \\ \partial \theta & [\hat{y}_0 \hat{\theta}^T] \end{bmatrix}. \tag{6}
\]

Remark 3. The analysis that follows may largely be repeated for other choices of estimator/optimizer combinations such as described in [12].

III. STABILITY ANALYSIS

Superficially, the proposed ES scheme might be seen as an example of the model-based ES framework discussed in [12]. However, the result in [12] requires \( k \) and \( \Omega \) to be chosen sufficiently small: small \( k \) ensures the optimizer is slow compared to the estimator; and small \( \Omega \) ensures the estimator is slow compared to the plant dynamics. The tuning strategy proposed in this paper is based on keeping the former of these time-scale separations while reversing the latter. In other words, \( \Omega \) is tuned to be sufficiently large so as to achieve accelerated estimation and subsequently accelerated optimization.

The static map output can be captured by the following Fourier series expansion

\[
z = f(\bar{u} + a \sin(\Omega t))^T \theta = \sum_{n \in \mathbb{Z}} \hat{f}_n(\bar{u}, a)^T \theta e^{in\Omega t}, \tag{7a}
\]

where

\[
\hat{f}_n(\bar{u}, a) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\bar{u} + a \sin(\tau)) e^{-in\tau} d\tau. \tag{7b}
\]
The Fourier representation of \( z \) characterizes its periodic behaviour when \( \bar{u} \) is held constant. Since \( z \) is the input to the linear dynamics \( F_o(s) \), after applying linear systems theory, the steady-state response of the output filter for a given \( \bar{u} \) can be written as
\[
x^*(\bar{u}, a) = \sum_{n \in \mathbb{Z}} (in\Omega I - A)^{-1} B f_n(\bar{u}, a) T \theta e^{in\Omega t}.
\] (8)

By letting \( \dot{x} = x - x^* \), and substituting in (2a) and (2b) it follows that
\[
\begin{align*}
\dot{x} &= A\dot{x} - \frac{\partial x^*}{\partial \bar{u}} \bar{u}, \\
y &= C\dot{x} + \sum_{n \in \mathbb{Z}} F_o(in\Omega) f_n(\bar{u}, a)^T \theta e^{in\Omega t}.
\end{align*}
\] (9a)

Importantly (9a)-(9b) hold even when \( \bar{u} \) is time-varying. The output equation (9b) can be further expanded as
\[
y = C\dot{x} + F_o(0) f(\bar{u})^T + F_o(\infty) [f(u) - f(\bar{u})]^T \theta + \sum_{n \in \mathbb{Z} \setminus \{0\}} (F_o(in\Omega) - F_o(\infty)) f_n(\bar{u}, a)^T \theta e^{in\Omega t}.
\] (10)

Therefore, the output of the plant can be regarded as a summation of the following terms:
- The first two terms in equation (10) are “slow” terms. \( C\dot{x} \) is the transient response of the output filter which is considerably slower than the periodic dither, \( \sin(\Omega t) \), for sufficiently large \( \Omega \). \( f(\bar{u}) \theta \) is a function of the optimizer output, \( \bar{u} \), which can be made arbitrarily slow by choosing \( k \) to be small. The sum of these two terms will be estimated by \( \hat{y}_0 \).
- The third term is a zero-mean periodic signal. This is the “useful” part of the output that can be used to estimate the unknown parameters \( \theta \).
- The last summation can be made arbitrary small, since \( (F_o(in\Omega) - F_o(\infty)) \rightarrow 0 \) for sufficiently large \( \Omega \). This will introduce an error term in the estimation.

By defining the following error terms
\[
\begin{align*}
\dot{y}_0 &= \hat{y}_0 - C\dot{x} - F_o(0) f(\bar{u})^T \theta, \\
\dot{\theta} &= \dot{\bar{u}} - F_o(\infty) \theta, \\
\bar{u} &= \bar{u} - u^* ,
\end{align*}
\] (11a)

and substituting (10)-(11c) into (6) we can find the exact equation for the closed loop error system, which is given by
\[
\begin{align*}
\begin{bmatrix} \dot{y}_0 \\ \dot{\theta} \end{bmatrix} &= -\Omega A\theta \begin{bmatrix} y_0 \\ \dot{\bar{u}} \end{bmatrix} + g_\theta + h(\bar{u} + u^*, a, \Omega t), \\
\bar{u} &= -k\Omega F_o(\infty) f'(\bar{u} + u^*)^T \theta + g_u,
\end{align*}
\] (12a)

where
\[
\begin{align*}
g_x &= -\frac{\partial x^*}{\partial \bar{u}} \bar{u}, \\
A\theta(\bar{u}, a, \Omega t) &= \begin{bmatrix} f_1(f_1)^T \end{bmatrix}; f_1 = f(u) - f(\bar{u}), \\
g_\theta &= \left[-C\dot{x} - F_o(0) \frac{\partial f_o(\bar{u}, a)^T \theta}{\partial \bar{u}} \bar{u} \right]_{0:1\times n}^T,
\end{align*}
\] (13a)

\[
h(\bar{u}, a, \Omega t) = \Omega \left[ \frac{1}{f_1} \sum_{n \in \mathbb{Z} \setminus \{0\}} (F_o(in\Omega) - F_o(\infty)) f_n(\bar{u}, a)^T e^{in\Omega \theta} \right] \\
g_u = -k\Omega f'(\bar{u} + u^*)^T \dot{\bar{u}}.
\] (13e)

Remark 4. In the new error coordinates, the goal of the ES scheme is to drive \( \dot{\bar{u}} \) to a small neighborhood of \( 0 \). Note that if \( \dot{\bar{u}} \rightarrow 0 \), then \( \bar{u} \rightarrow F_o(\infty) \theta \). I.e. the parameter estimate is scaled by \( F_o(\infty) \). This is not a concern, since this scaling can effectively be absorbed into \( k \).

The appealing feature of the above error system is that the dynamics of the estimator and the optimizer can be made arbitrarily fast by choosing a higher dither frequency. Also as \( \Omega \rightarrow \infty \), then the error term \( h(\bar{u}, a, \Omega t) \rightarrow 0 \).

If the perturbation term \( h(\bar{u}, a, \Omega t) \) is ignored, then it can be shown that (12a)-(12c) has an equilibrium at the origin. Moreover, without interconnection terms \( g_x, g_\theta \) and \( g_u \), stability analysis of the remaining unperturbed isolated system is straightforward. The premise in this analysis is that, each isolated subsystem is asymptotically stable. For the isolated estimator, this requires the signal vector \( [f_1(\bar{u}, a, \Omega t)]^T \) in (12b) to fulfil a persistency of excitation criterion. According to Definition 3.4.1 in [15], if the signal vector \( f_1 \) satisfies the condition
\[
\frac{1}{2\pi} \int_{-\infty}^{+\infty} f_1 f_1^T(\bar{u}, a, \tau) d\tau \geq \alpha_0 I,
\] (14)

for some \( \alpha_0 > 0 \) and \( \forall t > t_0 \), it is “persistently exciting of level \( \alpha_0 \)”.

Assumption 4. For all \( \bar{u} \in B^l(v) \) and for some \( m \geq n \), \( [f_1(\bar{u} + u^*, a) \ f_2(\bar{u} + u^*, a) \ ... \ f_m(\bar{u} + u^*, a)] \) is full rank.

Remark 5. When the nonlinear map \( f(\cdot)^T \theta \) is a polynomial with \( \theta \) representing the unknown coefficients, Assumption 4 can be satisfied with \( m = n_\theta \). In other cases, the satisfaction of this assumption needs to be investigated.

Lemma 1. The signal vector \( [f_1(\bar{u}, a, \Omega t)]^T \) is persistently exciting iff Assumption 4 holds. In that case \( \alpha_0 \) is \( O(a^{2m}) \).

Proof. Only a sketch of the proof is provided here. Further details are omitted for brevity. It can be shown that for a sufficiently small \( a \), \( f_i \) is \( O(a^i) \). Substituting the Fourier series representation for \( f_i(\cdot) \) into equation (14), the integral will be a summation of rank-1 positive symmetric semi-definite matrices of the form \( f_i f_i^T \). If Assumption 4 holds, the aforementioned summation results in a full rank positive-definite matrix with the smallest eigenvalue of \( O(a^{2m}) \). It follows directly from (14) that if \( f_i(\cdot) \) is persistently exciting, then \( [f_1(\cdot)^T]^T \) is persistently exciting of the same level.

It is worthwhile to note that in most black-box approaches, the local gradient of \( f(u) \) is the only parameter to be identified. However in this scheme \( n_\theta \) parameters should be identified. According to Lemma 1, the first \( m \) harmonics of \( f(u) \) must be linearly independent to be able to identify \( n_\theta \) parameters.
Fig. 2. \( \hat{\theta}_1(t) \) and \( \hat{\theta}_2(t) \) trajectories in \( \Omega t \) time-scale. Simulations have been done for \( \Omega \) equal to 200 (grey) and 2000 (black) rad/s.

To simplify the presentation here, \( k \) and \( \Omega \) are selected as

\[
k' = k/a^{2m}, \quad \Omega' = \Omega a^{2m}. \tag{15}
\]

The following theorem is the main result of this paper:

**Theorem 1.** Consider the system described by (1)-(2b) under Assumptions 1-4. For any given \((r_x, r_y, r_{\theta}, r_u) \) there exist \( a^*, k^*, \) and \( \Omega^* \) such that for all \((a, k', \Omega') \) \((0, a^*) \times (0, k^*) \times (\Omega^*, \infty)\), any trajectory \((\hat{x}(t), \hat{y}_0(t), \hat{\theta}(t), \hat{u}(t))\) originating in \( B^a(r_x) \times B^k(r_y) \times B^{\Omega}(r_\theta) \times B^1(v) \) will satisfy,

\[
\lim_{t \to \infty} \sup \|\hat{x}(t), \hat{y}_0(t), \hat{\theta}(t), \hat{u}(t)\| = O(a) \tag{16}
\]

**Proof:** See Appendix A.

**Remark 6.** This theorem has stronger assumptions than required in [12], but leads to the result that the rate of convergence (which is of interest in fast ES) is delivered through the use of large frequency dither. In Theorem 1, \( a, \) \( k', \) and \( \Omega' \) can be tuned independently. According to equation (12c), increasing \( \Omega' \) allows accelerated optimization.

**Remark 7.** By reducing the dither amplitude, it is possible to achieve convergence of \( \hat{u} \) to a smaller neighborhood of \( u^* \). This also requires \( k \) and \( \Omega \) to be chosen accordingly. This is an important factor to be considered by practitioners.

**Remark 8.** Although Theorem 1 is limited to bi-proper plant dynamics, it can be generalized to include proper plants as well. By using the results in [8], if the relative degree of the plant dynamics are known to the designer, appropriate output filters can be designed to achieve the same result.

**IV. SIMULATION EXAMPLE**

The following example is presented to show the performance of the proposed scheme. Consider a Hammerstein plant with \( z = \theta_2 u^2 + \theta_1 u \), and bi-proper plant dynamics, \( F_o(s) = (s + 2)/(s + 1) \). The unknown parameter vector \( \theta = [\theta_2, \theta_1] \) is equal to [1, 2], and \( u^* = -1 \). The ES scheme is tuned with \( a = 0.7, k = 0.001 \) and initialized with \( x(0) = 0, \theta(0) = [-1; -1], \) and \( \bar{u}(0) = 1 \).

![Fig. 2. \( \hat{\theta}_1(t) \) and \( \hat{\theta}_2(t) \) trajectories in \( \Omega t \) time-scale. Simulations have been done for \( \Omega \) equal to 200 (grey) and 2000 (black) rad/s.](image1)

![Fig. 3. \( y \) (solid) and \( f(u)^\top \theta \) (dashed) versus \( u \). Plant with bi-proper dynamics, \( F_o(s) = s + 2/s + 1 \). State of the plant shown at the initial condition (\( \circ \)), after 1s (\( \odot \)) and 50s (\( \square \)) and at the optimal point (\( * \)).](image2)

Fig. 2 shows the trajectories for \( \hat{\theta}_1(t) \) and \( \hat{\theta}_2(t) \) for two different dither frequencies in the \( \Omega t \) time-scale. The simulation durations were set at 5000\( \Omega^{-1} \) to provide the same duration for all frequencies in the \( \Omega t \) time-scale. For high frequencies the trajectories for the parameter estimator outputs follow almost the same trend. The same results can be shown for other signals in the closed loop.

Fig. 3 shows the plant output, \( y \), versus the optimizer output, \( \bar{u} \). The simulations for this part have been carried out with a fixed duration of 2000s only for two frequencies. For the small frequency, \( \Omega = 1 \text{ rad/s} \), the optimizer is slow compared to the plant dynamics. Therefore, the system output, \( y \), remains close to \( f(\bar{u})^\top \theta \). For the large frequency, the optimizer converges to \( u^* \) much faster than the system dynamics can settle to \( f(u^*)^\top \theta \). Also the state of the system is shown at \( t = 1 \text{ s} \) for \( \Omega = 2000 \text{ rad/s} \) case and \( t = 50 \text{ s} \) for \( \Omega = 1 \text{ rad/s} \) case.

**V. CONCLUSION**

This paper has proposed a fast model-based extremum seeking for a Hammerstein plant, where only the structure of the nonlinear map and the relative degree of the output filter need to be known. The proposed tuning strategy allows for accelerated estimation and optimization. The result is limited to SISO Hammerstein plants. Future work will include Wiener-Hammerstein plants, multi-input plants, and inclusion of noise in analysis.

**REFERENCES**


In Appendix B it is shown that there exists a Lyapunov function $V_\theta(t, \theta)$, such that

$$c_1\|\dot{\theta}\|^2 \leq V_\theta(t, \theta) \leq c_2\|\dot{\theta}\|^2,$$  \hspace{1cm} (17a)

$$\frac{\partial V_\theta}{\partial t} + \frac{\partial V_\theta}{\partial \theta} \dot{\theta} \leq -c_3 a^{2m}\Omega\|\dot{\theta}\|^2,$$  \hspace{1cm} (17b)

$$\|\dot{\theta}\|^2 \leq c_4\|\dot{\theta}\|^2,$$  \hspace{1cm} (17c)

where $c_1$, $c_2$, $c_3$ and $c_4$ are $O(1)$ positive constants. Since there exists such a $V_\theta$, one can use $\sqrt{V_\theta}$ instead.

The rate of change of the Lyapunov functions $V_x$, $V_u$, and $V_\nu$ in the presence of the interconnection terms can now be studied. This problem can be approached by doing worst case analysis in which interconnection effects are investigated using their upper bounds. It can then be shown that for sufficiently small $k'$ and sufficiently large $\Omega'$,

$$\frac{d}{dt} \left[ \frac{V_x}{\sqrt{V_\theta}} \right] \leq A_s(k') \left[ \frac{\|\dot{x}\|}{\Omega'} \left| \frac{\dot{\theta}}{\dot{\theta}} \right| \Omega' \|f'(\dot{\theta} + u_\ast)^T \theta\| \right],$$  \hspace{1cm} (18)

where

$$A_s(k') = \begin{bmatrix}
-c_{xx} & -c_{x\theta} & c_{xu} \\
-c_{ox} & 0 & c_{o\theta} & c_{o\nu}
\end{bmatrix},$$

and $c_{xx}, c_{x\theta}, ..., c_{uu}$ are all positive constants that are independent of $a$, $k'$, and $\Omega'$. Now let the Lyapunov function for the interconnected system be

$$V = [d_x \ d_\theta \ d_u]^T \sqrt{V_\theta} V_u^T,$$  \hspace{1cm} (19)

where $(d_x, d_\theta) \in \mathbb{R}_{>0}^2$ are yet to be determined. From (18), it is possible to ensure that $dV/dt$ is negative definite by picking $k'/d_\theta$, $d_\theta/d_x$, and $d_x$ to be sufficiently small. Therefore, $V$ will decrease in time towards zero, guaranteeing convergence of $(\dot{x}, \dot{\theta}, \dot{u})$ to the origin as long as the system states remain within the region for which (18) is valid. The weights, $d_x$, $d_\theta$ can be made arbitrary small (although this will mean $k'$ also has to be made small). Thus, the Lyapunov function can be made arbitrary close to $|\dot{u}|$. Since at $t = 0$, $|\dot{u}| < v$, then $V$ must increase if $\dot{u}$ is to leave $B^3(v)$. Thus, convergence to the origin of the unperturbed system is guaranteed.

B. Perturbation Effect

For a given dither amplitude, by picking $k'$ sufficiently small, the unperturbed-interconnected error system is asymptotically stable uniformly in $t$. However, the origin is not an equilibrium of the perturbed error system. In this case, stability analysis of the perturbed system can not be performed in the usual manner. Instead, it can be shown that the error states would be ultimately bounded by a small quantity, if the perturbation can be shown to be small in some sense.

The rate of change of $V$ along the trajectories of the error system (12a)-(12c) satisfies

$$\dot{V}(t, z) \leq -c_x\|\dot{x}\|^2 - c_\theta\|\dot{\theta}\|^2 - c_u k'\|f'(\dot{u} + u_\ast)^T \theta\| + c_4 d_\theta h(\dot{u} + u_\ast, a, \Omega t),$$  \hspace{1cm} (20)
where  
\[  c_{x} = (-c_{x}d_{x} + c_{x}d_{0}) , \]
\[  c_{\theta} = (k'c_{\theta}d_{x} - (c'_{\theta} - k'c_{\theta}d_{0})d_{0} + k'c_{\theta}d_{u}) , \]
\[  c_{u} = (c_{u}d_{x} + c_{u}d_{0} - c_{u}u) , \]
\[  c'_{4} = c_{4}/(2\sqrt{c_{1}}) , \]

and \( z \) represents \([\tilde{x}, \tilde{\theta}, \tilde{u}]\). The right-hand side of (20) is not always negative definite because, as trajectories approach the origin, the perturbation effect can make \( V \) positive. Using the ultimate boundedness notion [16], it is necessary to find a region \( \Lambda = \{ \dot{e} \leq V(z) \} \) for some \( e > 0 \) in which

\[  V(t, z) \leq -W(z) , \quad \forall z \in \Lambda, \forall t \geq t_{0} , \quad (21) \]

for some continuous positive definite function \( W(z) \). In fact, in this region, the above inequality is satisfied and therefore the trajectories behave as if the origin is asymptotically stable. According to (20) the region \( \Lambda \) can be specified by examining each of the following inequalities which satisfy (21) separately:

\[  -c_{x}\|\tilde{x}\| + c'_{4}d_{0}h(\tilde{u}, a, \Omega t) < 0 \]
\[  \Rightarrow \|\tilde{x}\| > \frac{c'_{4}d_{0}}{c_{x}}h(\tilde{u}, a, \Omega t) \quad (22a) \]

\[  -c_{\theta}\|\tilde{\theta}\| + c'_{4}d_{0}h(\tilde{u}, a, \Omega t) < 0 \]
\[  \Rightarrow \|\tilde{\theta}\| > \frac{c'_{4}d_{0}}{c_{\theta}\Omega}h(\tilde{u}, a, \Omega t) , \quad (22b) \]

\[  -c_{u}k'f'(\tilde{u} + u_{x})\|\tilde{\theta}\| + c'_{4}d_{0}h(\tilde{u}, a, \Omega t) < 0 \]
\[  \Rightarrow \|f'(\tilde{u} + u_{x})\| > \frac{c'_{4}d_{0}}{c_{u}k'\Omega}h(\tilde{u}, a, \Omega t) . \quad (22c) \]

The last inequality can be interpreted in terms of \( |\tilde{u}| \) based on Assumption 3. Therefore, the union of the sets (22a)-(22c) specify the region \( \Lambda \) where \( V < 0 \). Now consider the task of finding a value of \( \epsilon \) consistent with the definition of \( \Lambda \). To find this \( \epsilon \), first it is useful to note that

\[  V_{\min} \leq V \leq V_{\max} , \quad (23) \]

where

\[  V_{\min} = [d_{x} d_{0} 1][\lambda_{\min}^{1/2}(P)]\|\tilde{x}\|^{2} \sqrt{c_{1}}\|\tilde{\theta}\| \|\tilde{u}\|^{T} , \]
\[  V_{\max} = [d_{x} d_{0} 1][\lambda_{\max}^{1/2}(P)]\|\tilde{x}\|^{2} \sqrt{c_{2}}\|\tilde{\theta}\| \|\tilde{u}\|^{T} . \]

By using (22a)-(22c) and the RHS of (23) inequality it can be concluded that

\[  \epsilon = [d_{x} d_{0} 1][\lambda_{\min}^{1/2}(P)]\|\tilde{x}\|^{2} \sqrt{c_{1}}\|\tilde{\theta}\| \|\tilde{u}\|^{T} \leq \epsilon . \quad (24) \]

It can be shown that for sufficiently large dither frequency, \( h(\tilde{u}, a, \Omega t) \) is upper bounded by some quantity which is \( O(a) \). Therefore, from (24) it follows that \( \epsilon \) is \( O(a) \). That means, by decreasing \( a \), convergence to a smaller Lyapunov surface is guaranteed.

To find the ultimate bound on the error terms, the left inequality in (23) can be used to show that

\[  [d_{x} d_{0} 1][\lambda_{\min}^{1/2}(P)]\|\tilde{x}\|^{2} \sqrt{c_{1}}\|\tilde{\theta}\| \|\tilde{u}\|^{T} \leq \epsilon . \quad (25) \]

Since \( \epsilon \) is \( O(a) \), so is the LHS. It can be concluded that for sufficiently small dither amplitude, \( \|\tilde{x}\|, \|\tilde{\theta}\| \) and \( |\tilde{u}| \) converge to a ball with a \( O(a) \) radius.

APPENDIX B. LYAPUNOV FUNCTION FOR ISOLATED ESTIMATOR

According to Theorem 3.6.1 in [15], the isolated subsystem in (12b) (i.e. with \( g_{\theta} = h(\cdot) = 0 \)) is GES if Assumption 4 holds. In order to show exponential stability the Lyapunov function \( V_{e} = \|\bar{y}_{0} \dot{\bar{y}}^{T}T\|^{2}_{2} \) can be used. The time derivative of \( V_{e} \) for the isolated estimator dynamics in the \( \Omega t \) time-scale is

\[  V_{e} = \bar{y}_{0} \dot{\bar{y}} + \Omega \dot{\bar{y}} = -\left(1 f_{1}(\bar{u}, a, \tau)\right)^{T} \left[\bar{y}_{0} \dot{\bar{y}}\right]^{2} , \quad (26) \]

where \( f_{1} \) is a zero mean periodic signal vector. Under the persistency of excitation condition considered in Assumption 4 it can be shown that

\[  V_{e}(\bar{y}_{0}, \dot{\bar{y}}) \leq V_{e}(\bar{y}_{0}(t_{0}), \dot{\bar{y}}(t_{0})) e^{-b(t-t_{0})} , \quad (27a) \]

\[  \|\bar{y}_{0}(t) \dot{\bar{y}}(t)T\|^{2}_{2} \leq \|\bar{y}_{0}(t) \dot{\bar{y}}(t)T\|^{2}_{2} e^{-b(t-t_{0})} , \quad (27b) \]

in which

\[  \gamma = \frac{2\alpha_{0}T}{2 + \beta^{2}T^{2}} , \quad b = \frac{1}{T} \ln \frac{1}{1 - \gamma} , \quad (28) \]

\[  T = 2\pi/\Omega , \quad \beta = \sup_{t \geq 0} \|1 F^{T}T\|^{2} . \]

By using a Fourier series expansion of \( f_{1} \), it can be shown that, for sufficiently small dither amplitude, \( \|f_{1}\| \) is \( O(a) \) and therefore, \( \beta \rightarrow 1 \) as \( a \rightarrow 0 \).

By taking the square root of (27b), exponential convergence of \( \|\bar{y}_{0} \dot{\bar{y}}T\|^{2} \) can be established as

\[  \|\bar{y}_{0}(t) \dot{\bar{y}}(t)T\|^{2} \leq \rho(\|\bar{y}_{0}(t) \dot{\bar{y}}(t)T\|^{2} e^{-\lambda(t-t_{0})} \quad (29) \]

where \( \rho = 1/\sqrt{T-\gamma} \) and \( \lambda = b/2 \). \( \lambda \) is an important parameter since it determines how fast parameters converge. For a sufficiently small dither amplitude

\[  \lambda = \frac{1}{2T} \ln \frac{1}{1 - \gamma} = \frac{1}{2T} ln(1 + \frac{\gamma}{1 - \gamma}) = O(\alpha T^{2}) . \quad (30) \]

Similar analysis can reveal that \( \rho \rightarrow 1 \) as \( a \rightarrow 0 \).

Since \( V_{e} \) is negative semi-definite it can not be used in the interconnected system analysis. Letting \( \varphi(\sigma; t, \theta) \) denote the solution of the isolated estimator states at time \( \sigma \) with initial conditions \((t, \theta)\), the following Lyapunov function can be proposed as

\[  V_{\theta}(t, \theta) = \int_{t}^{t+\delta} \varphi^{T}(\sigma; t, \theta) \varphi(\sigma; t, \theta) d\sigma , \quad (31) \]

where \( \delta' \) is a positive \( O(1) \) constant. Due to the exponentially decaying bound on the trajectories established in (29) and based on the Converse Lyapunov theorem [16], it can be shown that such a Lyapunov function has the following properties:

\[  c_{1}\|\tilde{\theta}\|^{2} \leq V_{\theta}(t, \theta) \leq c_{2}\|\tilde{\theta}\|^{2} \quad (32a) \]

\[  \frac{\partial V_{\theta}}{\partial t} + \frac{\partial V_{\theta}}{\partial \theta} f(t, \theta) \leq -c_{3} a^{2m}\|\tilde{\theta}\|^{2} \quad (32b) \]

\[  \left| \frac{\partial V_{\theta}}{\partial \theta} \right| \leq c_{4}\|\tilde{\theta}\|^{2} \quad (32c) \]

where \( c_{1}, c_{2}, c_{3}, \) and \( c_{4} \) are positive \( O(1) \) quantities.