ADAPTIVE OBSERVER FOR DISCRETE TIME LINEAR TIME VARYING SYSTEMS

Arnaud Guyader * **and Qinghua Zhang** **

* *Université de Haute Bretagne, 35043 Rennes Cedex, France*
aguyader@irisa.fr
** IRISA-INRIA, *Campus de Beaulieu, 35042 Rennes Cedex, France*
zhang@irisa.fr

Abstract: For joint state-parameter estimation in discrete time stochastic multiple-input multiple-output linear time varying systems, an efficient adaptive observer is proposed in this paper. In the noise-free case, the global exponential convergence of the adaptive observer is first established. It is then proved that, in the noise-corrupted case, the state and parameter estimation errors remain bounded if the noises are bounded, and moreover, the estimation errors converge in the mean to zero if the noises have zero means.

Keywords: state and parameter estimation, discrete time system, linear time varying system.

1. INTRODUCTION

The Luenberger observer and the Kalman filter are well known solutions for state estimation in linear dynamic systems, in continuous time as well as in discrete time. For joint estimation of state and unknown parameters, some results are also known under the name of *adaptive observer*, see, *e.g.*, (Kreisselmeier, 1977; Bastin and Gevers, 1988; Marino and Tomei, 1995; Besançon, 2000; Zhang, 2002). These adaptive observers have been known *in continuous time*, and there are relatively few results of similar nature for *discrete time* systems. For single-input single-output (SISO) time invariant discrete time systems, some results can be found in (Landau, 1979; Ioannou and Kokotovic, 1983).

An adaptive observer is proposed in this paper for joint estimation of state and parameters in discrete time stochastic multiple-input multiple-output (MIMO) linear *time varying* systems. It is a discrete time counterpart of the continuous time algorithm presented in (Zhang, 2002). As seen in the formulation of the discrete time persistent excitation condition and in the convergence analysis, this adaptation from continuous time to discrete time is not trivial.

Let us consider discrete time stochastic MIMO linear time varying systems of the form

$$\theta_{k+1} = \theta_k + e_k \quad (1a)$$
$$x_{k+1} = A_k x_k + B_k u_k + \Psi_k \theta_k + w_k \quad (1b)$$
$$y_k = C_k x_k + v_k \quad (1c)$$

where $x_k \in \mathbb{R}^n, u_k \in \mathbb{R}^l, y_k \in \mathbb{R}^m$ are respectively the state, input, output of the system, A_k, B_k, C_k are known time varying matrices of appropriate sizes, $\theta_k \in \mathbb{R}^p$ is an unknown parameter vector, $\Psi_k \in \mathbb{R}^n \times \mathbb{R}^p$ is a matrix of known signals, and e_k, w_k, v_k are noises of appropriate dimensions. The time varying matrices A_k, B_k, C_k, Ψ_k are all assumed bounded. The noises e_k, w_k, v_k are bounded and have zero mean. Note that *no* whiteness of the noises is required in this paper.

The purpose of this paper is to design a recursive algorithm for joint estimation of the state vector x_k and the parameter vector θ_k from the input u_k, the output y_k, the excitation Ψ_k and the system matrices A_k, B_k, C_k.

The study of such systems is mainly motivated by fault detection and isolation (FDI) for which the term $\Psi_k \theta_k$ models the faults to be detected and isolated. See (Xu and Zhang, 2002) for some related work in continuous time. Another motivation is adaptive control for which the term $\Psi_k \theta_k$ models some modeling uncertainties.

Remark 1. For the purpose of state and parameter estimation, it will not make any extra difficulty if, in (1b), the term $B_k u_k$ is replaced by any known nonlinear functions of u_k or of any other known variables. Such nonlinearities have been considered (in the continuous time case) in (Bastin and Gevers, 1988; Marino and Tomei, 1995). For presentation clearness, let us assume the linear inputs in this paper. □

Remark 2. In the proposed method, *no* particular form of the matrices A_k and C_k is required, whereas classical methods typically assume some (time invariant) canonical form of the two matrices. This feature is particularly important for time varying systems which would require some non trivial transformation to achieve a canonical form. □

A natural idea for joint state and parameter estimation is to apply the Kalman filter to the extended system obtained by appending the unknown parameters θ_k into the state vector. Note that, even in the case of constant matrices A, B, C, the extended system is typically time varying, since the matrix Ψ_k should sufficiently excite the system in order to estimate the unknown parameters. In general, it is not easy to guarantee the convergence of the Kalman filter for *time varying* systems. Application of classical results requires uniform complete observability (Jazwinski, 1970). In practice, it is difficult to check the uniform complete observability of the *extended* system. Therefore, the analysis of the Kalman filter applied to the extended system is not a trivial problem. In this paper, instead of assuming the observability of the extended system, the proposed method is essentially based on the observability of the matrix pair (A_k, C_k) and on some persistent excitation condition.

In section 2 we first establish the exponential convergence of the proposed adaptive observer in the noise-free case. The noise-corrupted case is considered in section 3 where the boundedness of state and parameter estimation errors and their convergence in the mean to zero are proved. Section 4 is devoted to a numerical example. Finally, some concluding remarks are drawn in section 5.

2. THE NOISE-FREE CASE

In this section, the proposed adaptive observer is described and the exponential convergence to zero of the estimation errors is established in the noise-free case. It will be the basis for the proofs in the noise-corrupted case presented in the next section.

Throughout the paper, the Euclidean norm is used for vectors and the spectral norm[1] is used for matrices.

Definition 1. The linear time varying system

$$\eta_{k+1} = F_k \eta_k$$

is said *exponentially stable* if there exist two constants $r > 0$ and $0 < q < 1$ such that

$$\left\| \prod_{i=k_0}^{k-1} F_i \right\| \leq r q^{k-k_0}$$

for all $k > k_0$. □

Clearly, this definition implies $\|\eta_k\| \leq r q^{k-k_0} \|\eta_{k_0}\|$.

Assumption 1. The time varying matrices A_k and C_k are such that there exists a bounded time varying matrix $K_k \in \mathbb{R}^n \times \mathbb{R}^m$ so that the linear time varying system

$$\eta_{k+1} = (A_k - K_k C_k)\eta_k$$

is exponentially stable. □

Note that this assumption is equivalent to say that, when the term $\Psi_k \theta_k$ and the noises are absent in system (1), an exponential observer can be designed for the estimation of the state x_k. It is known that, if the time varying matrix pair (A_k, C_k) is completely uniformly observable, the Kalman gain will fulfill the requirement (Jazwinski, 1970).

As an adaptation of the continuous time adaptive observer presented in (Zhang, 2002), the proposed discrete time adaptive observer is as follows.

$$\Upsilon_{k+1} = (A_k - K_k C_k)\Upsilon_k + \Psi_k \tag{2a}$$

$$\hat{\theta}_{k+1} = \hat{\theta}_k + \mu_k \Upsilon_k^T C_k^T (y_k - C_k \hat{x}_k) \tag{2b}$$

$$\hat{x}_{k+1} = A_k \hat{x}_k + B_k u_k + \Psi_k \hat{\theta}_k + K_k (y_k - C_k \hat{x}_k) + \Upsilon_{k+1}(\hat{\theta}_{k+1} - \hat{\theta}_k) \tag{2c}$$

where $\Upsilon_k \in \mathbb{R}^n \times \mathbb{R}^p$ is a matrix sequence obtained by linearly filtering Ψ_k, the vector sequences $\hat{x}_k$ and $\hat{\theta}_k$ are respectively the state and parameter estimates, $\mu_k > 0$ is a bounded scalar gain sequence satisfying the following assumption.

Assumption 2. The scalar gain sequence $\mu_k > 0$ is small enough so that

$$\|\sqrt{\mu_k} C_k \Upsilon_k\| \leq 1 \tag{3}$$

for all $k \geq 0$, where $\|\cdot\|$ denotes the matrix spectral norm.

[1] The spectral norm of a matrix is associated with the Euclidean norm and equal to the largest singular value of the matrix.

Like in system identification, an assumption on persistent excitation is required for parameter estimation.

Assumption 3. The matrix of signals Ψ_k is persistently exciting so that the matrix sequence Υ_k (obtained by linearly filtering Ψ_k through (2a)) and the gain sequence μ_k satisfy, for some constant $\alpha > 0$, integer $L > 0$ and for all $k \geq 0$, the following inequality

$$\frac{1}{L}\sum_{i=k}^{k+L-1} \mu_i \Upsilon_i^T C_i^T C_i \Upsilon_i \geq \alpha I \tag{4}$$

Remark 3. Typically each term in the sum of (4) is rank deficient, since the number of outputs m (the number of rows of C_i) is typically smaller than the number of parameters p (the number of columns of Υ_i). Nevertheless, if Υ_i for different i vary in different "directions", the average stated in (4) can be positive definite. The positive definiteness of the sum of some matrices is typically required as a persistent excitation condition in system identification. See, *e.g.*, (Aström and Wittenmark, 1989). □

The property of Algorithm (2) in the noise-free case is stated in the following theorem.

Theorem 1. If the noises are absent in system (1), that is, $e_k = 0, w_k = 0, v_k = 0$ for all $k \geq 0$, then, under Assumptions 1–3, Algorithm (2) is a global exponential adaptive observer, *i.e.*, the estimation errors $\hat{x}_k - x_k$ and $\hat{\theta}_k - \theta_k$ tend to zero exponentially fast when $k \to \infty$. □

The proof of this theorem requires the two following lemmas.

Lemma 1. If the linear time varying system $\eta_{k+1} = F_k \eta_k$, with $\eta_k \in \mathbb{R}^n, F_k \in \mathbb{R}^n \times \mathbb{R}^n$, is exponentially stable (see definition 1), then

(1) for any bounded sequence $g_k \in \mathbb{R}^n$, the sequence z_k defined by $z_{k+1} = F_k z_k + g_k$ is bounded;
(2) for any sequence g_k tending to zero exponentially fast, the sequence z_k defined as above tends also to zero exponentially fast. □

For the proof of the first part of this lemma, one can see (Freeman, 1965, page 168). The proof of the second part is a straightforward extension.

Lemma 2. Let $\phi_k \in \mathbb{R}^m \times \mathbb{R}^p$ be a matrix sequence such that its spectral norm $\|\phi_k\| \leq 1$ for all $k \geq 0$. If there exist a real constant $\alpha > 0$ and an integer $L > 0$ such that for all $k \geq 0$ the following inequality holds

$$\frac{1}{L}\sum_{i=k}^{k+L-1} \phi_i^T \phi_i \geq \alpha I \tag{5}$$

then the linear time varying system

$$z_{k+1} = (I - \phi_k^T \phi_k) z_k$$

is exponentially stable. □

A proof of this lemma can be found in Appendix A.

Now we are ready to prove Theorem 1.

Proof of Theorem 1. Define the error sequences

$$\tilde{x}_k = \hat{x}_k - x_k, \quad \tilde{\theta}_k = \hat{\theta}_k - \theta_k$$

In the absence of the noises in (1), following (1b) and (2c) it is easy to obtain the error dynamics

$$\begin{aligned}\tilde{x}_{k+1} &= A_k \tilde{x}_k + \Psi_k \tilde{\theta}_k + K_k(y_k - C_k \hat{x}_k) \\ &\quad + \Upsilon_{k+1}(\hat{\theta}_{k+1} - \hat{\theta}_k)\end{aligned}$$

According to (1a) and (1c), $\theta_{k+1} = \theta_k$ and $y_k = C_k x_k$ (it is assumed $e_k = 0, v_k = 0$), then

$$\begin{aligned}\tilde{x}_{k+1} &= (A_k - K_k C_k)\tilde{x}_k + \Psi_k \tilde{\theta}_k \\ &\quad + \Upsilon_{k+1}(\tilde{\theta}_{k+1} - \tilde{\theta}_k)\end{aligned} \tag{6}$$

The key step of the proof is to define the linearly combined error sequence

$$\eta_k = \tilde{x}_k - \Upsilon_k \tilde{\theta}_k \tag{7}$$

It is straightforward to compute the dynamic equation of η_k:

$$\begin{aligned}\eta_{k+1} &= (A_k - K_k C_k)\eta_k \\ &\quad + [(A_k - K_k C_k)\Upsilon_k + \Psi_k - \Upsilon_{k+1}]\tilde{\theta}_k\end{aligned}$$

Because Υ_k is generated from (2a), the last equation simply becomes

$$\eta_{k+1} = (A_k - K_k C_k)\eta_k$$

According to Assumption 1, the sequence η_k tends to zero exponentially fast.

Now let us study the error $\tilde{\theta}_k = \hat{\theta}_k - \theta_k$. Following (2b) and (1a), (1c) with $e_k = 0$ and $v_k = 0$,

$$\tilde{\theta}_{k+1} = \tilde{\theta}_k - \mu_k \Upsilon_k^T C_k^T C_k \tilde{x}_k \tag{8}$$

Substitute $\tilde{x}_k$ with (7), then the error equation becomes

$$\begin{aligned}\tilde{\theta}_{k+1} &= \tilde{\theta}_k - \mu_k \Upsilon_k^T C_k^T C_k(\eta_k + \Upsilon_k \tilde{\theta}_k) \\ &= (I - \mu_k \Upsilon_k^T C_k^T C_k \Upsilon_k)\tilde{\theta}_k \\ &\quad - \mu_k \Upsilon_k^T C_k^T C_k \eta_k\end{aligned} \tag{9}$$

According to Assumptions 2, 3 and Lemma 2, the homogeneous part of (9), that is, the linear time varying system

$$z_{k+1} = (I - \mu_k \Upsilon_k^T C_k^T C_k \Upsilon_k) z_k \tag{10}$$

is exponentially stable.

The sequences μ_k, C_k have been assumed bounded, and the boundedness of Υ_k is a consequence of the boundedness of Ψ_k and of Assumption 1, following Lemma 1.

Then following Lemma 1, the sequence $\tilde{\theta}_k$ driven by the exponentially vanishing sequence

$$-\mu_k \Upsilon_k^T C_k^T C_k \eta_k$$

through (9) tends to zero exponentially fast.

Finally, $\tilde{x}_k = \eta_k + \Upsilon_k \tilde{\theta}_k$ tends also to zero exponentially fast. □

Remark 4. It is clear in the proof of Theorem 1 that Assumptions 2 and 3 are for the purpose of ensuring the exponential stability of the linear time varying system (10). There are other ways to ensure this stability. A direct condition is

$$\left\| \prod_{i=k}^{k+L-1} (I - \mu_k \Upsilon_k^T C_k^T C_k \Upsilon_k) \right\| \leq \gamma$$

for some $L > 0$, $0 < \gamma < 1$ and for all $k \geq 0$. The disadvantage of this condition is that it requires the computation of the matrix products. The condition stated with Assumptions 2 and 3 is probably not necessary, but it is sufficient and requires simple computations as stated in (3) and (4). □

3. THE NOISE-CORRUPTED CASE

Now let us study the properties of the same algorithm (2) applied to the noise-corrupted system (1).

Theorem 2. Under Assumptions 1–3, when algorithm (2) is applied to system (1) with bounded noises e_k, w_k, v_k, the estimation errors $\hat{x}_k - x_k$ and $\hat{\theta}_k - \theta_k$ remain bounded.

Moreover, if the noises e_k, w_k, v_k have zero mean, then the estimation errors tend to zero in the mean, that is, when $k \to \infty$, the mathematical expectations $\mathbf{E}(\hat{x}_k - x_k) \to 0$, $\mathbf{E}(\hat{\theta}_k - \theta_k) \to 0$, and the convergence is exponentially fast. □

Proof of Theorem 2. The proof of this theorem essentially relies on the result already established in the noise-free case. Like in the proof of Theorem 1, the equations of the errors $\tilde{x}_k = \hat{x}_k - x_k$ and $\tilde{\theta}_k = \hat{\theta}_k - \theta_k$ are first derived, but now the noises are involved:

$$\begin{aligned} \tilde{x}_{k+1} &= (A_k - K_k C_k)\tilde{x}_k + \Psi_k \tilde{\theta}_k \\ &\quad + \Upsilon_{k+1}(\tilde{\theta}_{k+1} - \tilde{\theta}_k) \\ &\quad - w_k + K_k v_k + \Upsilon_{k+1} e_k & \text{(11a)} \\ \tilde{\theta}_{k+1} &= \tilde{\theta}_k - \mu_k \Upsilon_k^T C_k^T C_k \tilde{x}_k + \mu_k \Upsilon_k^T C_k^T v_k & \text{(11b)} \end{aligned}$$

According to Theorem 1, the estimation errors tend to zero exponentially fast in the noise-free case. It means that the error system (11) is, in the absence of the terms involving the noises, exponentially stable (as a linear time varying system).

The sequences K_k, C_k are bounded by assumption, and Υ_k is bounded following the boundedness of Ψ_k. For bounded noises, the terms involving the noises in (11) are thus all bounded. Then according to Lemma 1, the errors $\tilde{x}_k$ and $\tilde{\theta}_k$ governed by (11) are also bounded.

Now let us take the mathematical expectation at both sides of (11). Notice that the sequences μ_k, C_k, K_k, Ψ_k are deterministic, and so is Υ_k. Then, following the zero mean assumption on the noises, equations (11) after the mathematical expectation becomes

$$\begin{aligned} \mathbf{E}\tilde{x}_{k+1} &= (A_k - K_k C_k)\mathbf{E}\tilde{x}_k + \Psi_k \mathbf{E}\tilde{\theta}_k \\ &\quad + \Upsilon_{k+1}(\mathbf{E}\tilde{\theta}_{k+1} - \mathbf{E}\tilde{\theta}_k) \\ \mathbf{E}\tilde{\theta}_{k+1} &= \mathbf{E}\tilde{\theta}_k - \mu_k \Upsilon_k^T C_k^T C_k \mathbf{E}\tilde{x}_k \end{aligned}$$

These equations are the same as the error equations (6) and (8) in the noise-free case, except that the errors $\tilde{x}_k$ and $\tilde{\theta}_k$ are respectively replaced by their mathematical expectation $\mathbf{E}\tilde{x}_k$ and $\mathbf{E}\tilde{\theta}_k$. Then following the same procedure as in the proof of Theorem 1, the mathematical expectations $\mathbf{E}\tilde{x}_k$ and $\mathbf{E}\tilde{\theta}_k$ tend to zero exponentially fast. □

4. NUMERICAL EXAMPLE

Let us illustrate the behavior of the proposed adaptive observer with the simulation of a controlled satellite. The classic linearized satellite model, in its continuous time version, can be found in (Brockett, 1970). The satellite nominal orbit is assumed to be circular with the radius normalized to 1. The nominal angular velocity of the satellite is 3.49×10^{-4}rad/s. The equations of motion of the satellite are linearized around the nominal orbit. In order to obtain a discrete time model, the linearized system is sampled with the period $T_s = 0.1$s and with zero-order holders applied at its inputs. The obtained discrete time model is

$$\begin{aligned} x_{k+1} &= \begin{bmatrix} 1 & 0.1 & 0 & 3.49\times10^{-6} \\ 3.66\times10^{-8} & 1 & 0 & 6.98\times10^{-5} \\ -4.25\times10^{-14} & -3.49\times10^{-6} & 1 & 0.1 \\ -1.28\times10^{-12} & -6.98\times10^{-5} & 0 & 1 \end{bmatrix} x_k \\ &+ \begin{bmatrix} 0.1 & 0.005 & 0 & 1.16\times10^{-7} \\ 1.83\times10^{-9} & 0.1 & 0 & 3.49\times10^{-6} \\ -1.06\times10^{-15} & -1.16\times10^{-7} & 0.1 & 0.005 \\ -4.25\times10^{-14} & -3.49\times10^{-6} & 0 & 0.1 \end{bmatrix} \\ &\cdot \begin{bmatrix} 0 & 0 \\ \theta^1 & 0 \\ 0 & 0 \\ 0 & \theta^2 \end{bmatrix} u_k \\ y_k &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} x_k \end{aligned}$$

where the components of the state vector $x_k \in \mathbb{R}^4$ correspond to radial position, radial velocity, angular position and angular velocity, the components of the input vector $u_k \in \mathbb{R}^2$ are the radial

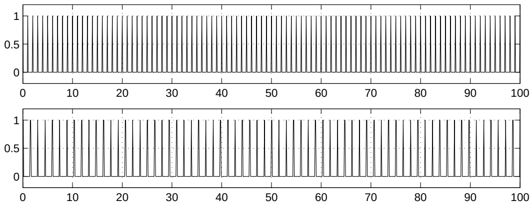

Figure 1. Input signals u_k^1 and u_k^2 used in the controlled satellite simulation. The time unit is second.

and tangential thrusts, the output vector $y_k \in \mathbb{R}^2$ correspond to distance and angle observations, and the constant coefficients θ^1 and θ^2 represent the efficiencies of the radial and tangential thrusts.

Note that the 4×4 matrix in the term involving u_k is due to the discretization of the continuous time model, since the coefficients θ^1 and θ^2 were originally defined in the continuous time model. In order to put the model into the form of (1), this term is reformulated as

$$\Psi_k \theta = H \begin{bmatrix} 0 & 0 \\ u_k^1 & 0 \\ 0 & 0 \\ 0 & u_k^2 \end{bmatrix} \begin{bmatrix} \theta^1 \\ \theta^2 \end{bmatrix}$$

where H is the aforementioned 4×4 matrix.

In the simulation, the parameter values are set to $\theta^1 = 1$ and $\theta^2 = 1.5$. The square impulse signals shown in figure 1 are used as inputs. The two simulated outputs are both disturbed by a Gaussian white noise whose standard deviation is 0.01.

The initial values used in the simulation are $x_0 = [1, 0, 0, 3.49 \times 10^{-4}]^T$, $\hat{x}_0 = [0.9, 0, 0, 3.14 \times 10^{-4}]^T$, $\hat{\theta}_0 = [0.5, 0.5]$. The adaptive observer parameters are

$$\mu_k \equiv 4, \qquad K_k \equiv \begin{bmatrix} 0.1412 & 4.93 \times 10^{-6} \\ 0.0932 & 5.26 \times 10^{-5} \\ -4.93 \times 10^{-6} & 0.1412 \\ -5.26 \times 10^{-5} & 0.0932 \end{bmatrix}$$

In figures 1–3 are, respectively, plotted the input signals, the state estimation errors, and the parameter estimates. Notice that, due to the noises added to the outputs y_k, the estimation errors randomly oscillate around zero instead of tending to zero. According to Theorem 2, the means of the estimation errors tend to zero when $k \to \infty$.

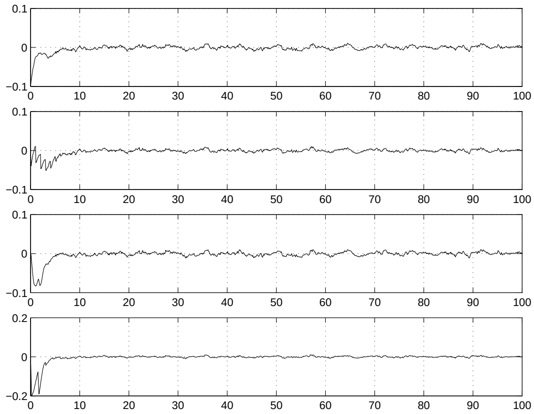

Figure 2. State estimation errors $\tilde{x}_k^1, \tilde{x}_k^2, \tilde{x}_k^3, \tilde{x}_k^4$. The time unit is second.

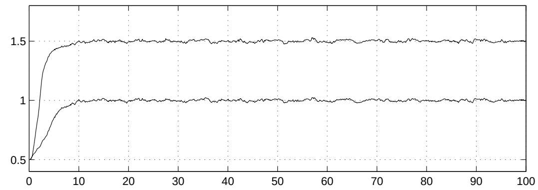

Figure 3. Parameter estimates $\hat{\theta}_k^1$ (lower) and $\hat{\theta}_k^2$ (upper). The true parameter values are $\theta^1 = 1$ and $\theta^2 = 1.5$. The time unit is second.

5. CONCLUSION

A numerically efficient adaptive observer has been proposed in this paper for joint state-parameter estimation in discrete time stochastic multiple-input multiple-output linear time varying systems. Essentially, if an exponential state observer can be designed for a linear system, then an adaptive observer can be designed for the system obtained after adding additive terms with unknown coefficients. A persistent excitation condition is required in order to ensure the convergence of the adaptive observer. The boundedness and convergence in the mean of the estimation errors have been proved under the assumption of bounded and zero mean noises. The analysis of the covariances of the estimation errors would require some assumption on the decay of the noises correlations and will be reported elsewhere. Potential applications of the proposed algorithm are fault detection and isolation, and adaptive control.

APPENDIX A. PROOF OF LEMMA 2

It is first observed that condition (5) is equivalent to

$$\frac{1}{L} \sum_{i=k}^{k+L-1} \|\phi_k w\| \geq \beta \tag{12}$$

for some positive constant β and for all unitary vector $w \in \mathbb{R}^p$. This equivalence is based on the following simple inequalities (for any unitary w)

$$w^T \phi_k^T \phi_k w \leq \|\phi_k^T \phi_k w\| \leq \phi_{\max} \|\phi_k w\|$$

with

$$\phi_{\max} = \sup_k \|\phi_k\|$$

and on the Cauchy-Schwarz inequality

$$\sum_{i=k}^{k+L-1} w^T \phi_k^T \phi_k w \geq \frac{1}{L}\left(\sum_{i=k}^{k+L-1} \|\phi_i w\|\right)^2$$

Consider the Lyapunov function candidate

$$V_k = z_k^T z_k$$

then

$$\begin{aligned} V_k - V_{k+1} &= z_k^T(\phi_k^T \phi_k) z_k + z_k^T[\phi_k^T \phi_k - (\phi_k^T \phi_k)^2] z_k \\ &\geq z_k^T(\phi_k^T \phi_k) z_k \end{aligned}$$

where the inequality is due to the assumption that the matrix spectral norm $\|\phi_k\| \leq 1$. Then

$$\begin{aligned} V_k - V_{k+L} &\geq \sum_{i=k}^{k+L-1} z_i^T(\phi_i^T \phi_i) z_i \\ &\geq \frac{1}{L}\left(\sum_{i=k}^{k+L-1} \|\phi_i z_i\|\right)^2 \end{aligned}$$

where the last inequality follows Cauchy-Schwarz.

Now a lower bound of $\sum_{i=k}^{k+L-1} \|\phi_i z_i\|$ is needed. Let us proceed as follows.

$$\sum_{i=k}^{k+L-1} \|\phi_i z_i\| \geq \sum_{i=k}^{k+L-1} \|\phi_i z_k\| - \sum_{i=k}^{k+L-1} \|\phi_i (z_k - z_i)\|$$

The first term at the right hand side is bounded from below according to (12):

$$\sum_{i=k}^{k+L-1} \|\phi_i z_k\| \geq \beta L \|z_k\|$$

For the second term,

$$\sum_{i=k}^{k+L-1} \|\phi_i(z_k - z_i)\| \leq \phi_{\max} L \sup_{k \leq i \leq k+L-1} \|z_k - z_i\|$$

It turns out that

$$\begin{aligned} \sup_{k \leq i \leq k+L-1} \|z_k - z_i\| &\leq \sum_{i=k}^{k+L-1} \|z_i - z_{i+1}\| \\ &= \sum_{i=k}^{k+L-1} \|\phi_i^T \phi_i z_i\| \\ &\leq \phi_{\max} \sum_{i=k}^{k+L-1} \|\phi_i z_i\| \end{aligned}$$

Therefore,

$$\sum_{i=k}^{k+L-1} \|\phi_i z_i\| \geq \beta L \|z_k\| - \phi_{\max}^2 L \sum_{i=k}^{k+L-1} \|\phi_i z_i\|$$

or equivalently,

$$\sum_{i=k}^{k+L-1} \|\phi_i z_i\| \geq \frac{\beta L}{1 + \phi_{\max}^2 L} \|z_k\|$$

It then follows

$$V_k - V_{k+L} \geq \frac{\beta^2 L}{(1 + \phi_{\max}^2 L)^2} \|z_k\|^2$$

The inequality (12) implies $\beta \leq \phi_{\max}$, then

$$0 < \gamma = \frac{\beta^2 L}{(1 + \phi_{\max}^2 T)^2} < 1$$

It follows that

$$V_{k+L} \leq (1 - \gamma) V_k$$

The exponential convergence to zero of $V_k = \|z_k\|^2$ is thus established.

6. REFERENCES

Aström, Karl J and Bjorn Wittenmark (1989). *Adaptive control.* Addison-Wesley. Reading, MA.

Bastin, Georges and Michel Gevers (1988). Stable adaptive observers for nonlinear time varying systems. *IEEE Trans. on Automatic Control* **33**(7), 650–658.

Besançon, Gildas (2000). Remarks on nonlinear adaptive observer design. *Systems and Control Letters* **41**(4), 271–280.

Brockett, Roger W. (1970). *Finite dimensional linear systems.* J. Wiley and sons. New York.

Freeman, Herbert (1965). *Discrete-time systems – an introduction to the theory.* John Wiley and Sons. New York.

Ioannou, P. A. and P. V. Kokotovic (1983). *Adaptative systems with reduced models.* Vol. 47 of *Lecture Notes in Control and Information Sciences.* Springer. Berlin.

Jazwinski, A. H. (1970). *Stochastic Processes and Filtering Theory.* Vol. 64 of *Mathematics in Science and Engineering.* Academic Press. New York.

Kreisselmeier, Gerhard (1977). Adaptive observers with exponential rate of convergence. *IEEE Trans. on Automatic Control* **22**(1), 2–8.

Landau, Ioan Doré (1979). *Adaptive control: the model reference approach.* M. Dekker. New York.

Marino, Riccardo and Patrizio Tomei (1995). *Nonlinear control design.* Information and system sciences. Prentice Hall. London, New York.

Xu, Aiping and Qinghua Zhang (2002). Fault detection and isolation based on adaptive observers for linear time varying systems. In: *IFAC World Congress.* Barcelona.

Zhang, Qinghua (2002). Adaptive observer for multiple-input-multiple-output (MIMO) linear time varying systems. *IEEE Trans. on Automatic Control* **47**(3), 525–529.