Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Revisiting different adaptive observers through a unified formulation

Qinghua ZHANG
INRIA-IRISA, Campus de Beaulieu, 35042 Rennes Cedex, France
Email: zhang@irisa.fr

***Abstract*— An adaptive observer is a recursive algorithm for joint estimation of states and parameters in dynamic systems. In this paper, 4 adaptive observers published by different authors are revisited through a unified formulation. Though these algorithms are apparently quite different as presented in the original publications, typically in their design procedures they use both static and dynamic transformations which can be reformulated in a unified manner. This new perspective allows to better understand each algorithm and helps to develop new algorithm variants for particular purposes.**

***Index Terms*— adaptive observer, state and parameter estimation, dynamic transformation.**

I. INTRODUCTION

An adaptive observer is a recursive algorithm for joint state-parameter estimation in dynamic systems, or for state estimation only despite the presence of unknown parameters. Such algorithms have important applications in fault detection and isolation (FDI) and in adaptive control. Some early works on adaptive observers for linear systems can be found in [1], [2]. Adaptive observers for nonlinear systems then have drawn more attentions of researchers. The algorithms presented in [3], [4] assume that the considered nonlinear systems can be linearized by coordinate change and output injection. The results reported in [5], [6], [7] do not require linearization, instead, they assume the existence of some Lyapunov function satisfying particular conditions. Recently, a constructive algorithm has been proposed in [8] for a particular class of nonlinear systems, with no resort to linearization.

A natural idea for joint estimation of states and parameters is to consider the extended system obtained by appending the unknown parameters into the state vector, and then to apply the Kalman filter or similar algorithms to the extended system. See, for example, [9], [10]. A discussion about the relation between such algorithms and adaptive observers is given in the appendix of this paper.

The purpose of this paper is to study the relationship between several existing adaptive observers through a general formulation. This study helps to better understand each of these existing algorithms with new insights and also to derive new variants. Most of these algorithms have been originally presented in papers dealing with nonlinear systems, however, the core of each adaptive algorithm is related to a linear system, typically obtained by coordinate change and output injection. In this paper, for simplicity of presentation, the problem of adaptive estimation will be directly formulated for linear systems. The study of such linear algorithms is also useful for developing adaptive observers for nonlinear systems which cannot be linearized. See [11], [8] for such examples.

The paper is organized as follows. In section II, the common steps used in different adaptive observers are first formulated. In section III, 4 different algorithms are revisited through a unified framework. Section IV concludes the paper.

II. COMMON STEPS IN THE DESIGN OF ADAPTIVE OBSERVERS

As mentioned in the introduction, in this paper the design of adaptive observers is directly considered for linear state space systems, though they could be linearized nonlinear systems as described in some references revisited in this paper. Let us first formulate the general state space systems considered in this paper:

$$\dot{\eta}(t) = A(t)\eta(t) + B(t)u(t) + \Phi(t)\theta \quad (1a)$$

$$y(t) = C(t)\eta(t) \quad (1b)$$

where $\eta(t) \in \mathbb{R}^n, u(t) \in \mathbb{R}^l$ and $y(t) \in \mathbb{R}^m$ are respectively the state, input and output of the system, $A(t), B(t), C(t)$ are known time varying matrices of appropriate sizes, $\theta \in \mathbb{R}^p$ is an *unknown* constant parameter vector, $\Phi(t) \in \mathbb{R}^{n\times p}$ is a matrix of known signals. The variables in $A(t), B(t), C(t), \Phi(t)$ are all assumed piecewise continuous and bounded in time. An adaptive observer should jointly estimate $\eta(t)$ and θ from measured $y(t), u(t)$ and $\Phi(t)$. Notice that appending θ into the state vector would yield a linear system to which the Kalman filter can be applied. See the appendix of this paper for a discussion about this approach.

For adaptive control, the term $\Phi(t)\theta$ can be used to deal with modeling uncertainties, whereas for FDI, it can be used to model faults to be detected and isolated.

The formulation (1) includes general linear time varying (LTV) multi-input multi-output (MIMO) systems. Some of the algorithms revisited in this paper consider a particular case (time invariant and single output) of such systems.

Remark that this class of systems can be extended to the so-called *state-affine nonlinear* systems, by replacing the time varying matrices $A(t), B(t), C(t), \Phi(t)$ with matrices depending also on $u(t)$ and $y(t)$. For example, $B(t)$ would be written as $B(t, u(t), y(t))$. The revisited adaptive observers can also be applied to such systems. However, for presentation simplicity, let us stay with the model formulated in (1).

Four adaptive observers developed by different authors [3], [4], [7], [12] will be revisited in this paper. Though they have been presented in quite different ways in the original references, typically they need two transformations simplifying the adaptive estimation problem. The first one aims at transforming the matrix pair $A(t), C(t)$ to some special form, and the second one has the purpose of simplifying the presence of θ in system equations.

A. Static transformation

The first transformation is static and is simply based on some coordinate change and/or on output injection. Let $T \in \mathbb{R}^{n\times n}$ be an invertible matrix and apply the corresponding coordinate change $x(t) = T\eta(t)$ to system (1) with output injection. The transformed system is then written as

$$\dot{x}(t) = A_o(t)x(t) + \varphi(t) + \Psi(t)\theta \tag{2a}$$

$$y(t) = C_o(t)x(t) \tag{2b}$$

with

$$\begin{aligned} C_o(t) &= C(t)T^{-1} \\ A_o(t) &= TA(t)T^{-1} - K(t)C_o(t) \\ \Psi(t) &= T\Phi(t) \\ \varphi(t) &= TB(t)u(t) + K(t)y(t) \end{aligned}$$

where $K(t) \in \mathbb{R}^{n\times m}$ is a chosen matrix for output injection. Notice that the coordinate change has been defined by a constant matrix T. One of the adaptive observers revisited in this paper does need a time varying coordinate change. But this detail is not essential at this stage and let us assume a constant T in order to simplify the presentation.

If the output injection is not used, then $A_o(t) = TA(t)T^{-1}$. It is well known that the coordinate change does not change the dynamic behavior (to be more clear, in the constant case, A_o and A have the same eigenvalues). The purpose of output injection is to modify the dynamic behavior of $A_o(t)$.

The above transformation is used in adaptive observer design procedures in order to put the matrix pair $A_o(t), C_o(t)$ into some desired form. The actual form depends on each particular adaptive observer and will be specified when the algorithm is revisited.

B. Dynamic transformation

Starting from a system in the form of (2), typically a dynamic transformation is further performed in adaptive observer design procedures. Though the transformations used by the revisited algorithms are apparently quite different, they can all be unified with a general formula. As detailed later, each of the revisited algorithms corresponds to some particular choices in the following transformation. Let $z(t) \in \mathbb{R}^n$ be the new transformed state vector, the general dynamic transformation takes the form,

$$\dot{\Xi}(t) = F\,\Xi(t) + G\Psi(t) \tag{3a}$$

$$\Omega(t) = H\,\Xi(t) \tag{3b}$$

$$z(t) = x(t) - \Omega(t)\theta \tag{3c}$$

where F, G, H are matrices of appropriate sizes defining the dynamic transformation, $\Xi(t)$ and $\Omega(t)$ are both matrices of variables and can be understood as the "state" and the "output" of the "state space system" defined by (3a) and (3b). The sizes of the matrices $F, G, H, \Xi(t), \Omega(t)$ depend on the choice of each particular adaptive observer and will be specified in each case. When the considered system is time varying, a time varying matrix $F(t)$ will be used. In any case no time dependence of the matrices G, H is needed.

After application of this transformation to system (2), it is straightforward to derive

$$\begin{aligned} \dot{z}(t) &= A_o(t)z(t) + \varphi(t) \\ &\quad + [A_o(t)H\Xi(t) + \Psi(t) - HF\Xi(t) - HG\Psi(t)]\theta \end{aligned} \tag{4a}$$

$$y(t) = C_o(t)z(t) + C_o(t)\Omega(t)\theta \tag{4b}$$

In the new transformed system, the part $A_o(t)z(t) + \varphi(t)$ remains unchanged, whereas the appearance of θ varies depending on the choice of the transformation defined by F, G, H. Appropriate choices of F, G, H can simplify the presence of θ in the transformed system, as shown in the next section.

The adaptive observers revisited below are designed for the estimation of $z(t)$ and θ. The estimation of $x(t)$ and $\eta(t)$ can then be obtained through the equalities $x(t) = z(t) + \Omega(t)\theta$ and $\eta(t) = T^{-1}x(t)$.

III. REVISITING DIFFERENT ADAPTIVE OBSERVERS THROUGH THE GENERAL FORMULATION

Four adaptive observers published by different authors are reformulated in the following with the aid of the general transformations introduced in the previous section. Notice that the notations used here are different from those used in the original publications, in order to reformulate the different algorithms in a unified framework.

A. Algorithm 1

Let us first revisit the adaptive observer proposed in [3] where only single output systems with *constant matrices* A, C are considered. It is assumed in [3] that[1] the pair (A, C) is observable, and that after the static transformation the system corresponding to (2) has the particular form

$$\dot{x}(t) = A_o x(t) + \varphi(t) + \Psi(t)\theta \tag{5a}$$

$$y(t) = c_o x(t) \tag{5b}$$

[1]To be more complete, the authors of [3] originally consider nonlinear systems which are linearized through coordinate change and output injection. This note applies also to the reference [4] of Algorithm 2.

with

$$A_o = \begin{bmatrix} 0 & a_{12} \\ 0 & A_{22} \end{bmatrix}, \quad c_o = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$$

where a_{12} is a $1 \times (n-1)$ vector and A_{22} *is an asymptotically stable* $(n-1) \times (n-1)$ *matrix*.

A dynamic transformation of this system has been introduced in [3] in order to design an adaptive observer. This particular transformation can be reformulated in the form of (3) with the following choices in the transformation:

$$F = A_{22} \qquad G = \begin{bmatrix} 0 & 1 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \cdots & \ddots & 0 \\ 0 & 0 & \cdots & \cdots & 1 \end{bmatrix}_{(n-1)\times n}$$

$$H = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ 1 & \cdots & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & \ddots & 0 \\ 0 & \cdots & \cdots & 1 \end{bmatrix}_{n\times(n-1)}$$

where the size of F is $(n-1) \times (n-1)$, the sizes of G, H are as indicated above. Accordingly, the sizes of $\Xi(t)$ and $\Omega(t)$ in (3) are respectively $(n-1) \times p$ and $n \times p$.

With these particular choices of the transformation applied to (5), and after some straightforward computation, the transformed system (4) becomes

$$\dot{z}(t) = A_o z(t) + \varphi(t) + \gamma \xi^T(t)\theta \tag{6a}$$

$$y(t) = c_o z(t) \tag{6b}$$

where

$$\gamma = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad \xi^T(t) = a_{12}\Xi(t) + \Psi_1(t)$$

with Ψ_1 being the first row of Ψ.

Notice that, due to the particular value of the vector γ, only the first state equation in (6a) is affected by the scalar quantity $\xi^T(t)\theta$. In this sense, the presence of the unknown parameter θ in the transformed state equation is simplified. Notice that, due to the special form of c_o, the first state component $z_1(t)$ is directly observed by $y(t)$. The other components of $z(t)$ can be easily estimated through the last $n-1$ state equations in (6a) due to the asymptotic stability of A_{22}. If the derivative $\dot{y}(t)$ is also considered as known, then the only unknown in the first state equation is θ. Therefore, intuitively, now it is easy to estimate θ. Of course, the adaptive observer as derived in [3] does not require the numerical differentiation of $y(t)$. It is formulated in the form of ordinary differential equations (ODE):

$$\begin{aligned} \dot{\hat{z}}(t) &= A_o\hat{z}(t) + \varphi(t) + \gamma\xi^T(t)\hat{\theta}(t) \\ &\quad + k[y(t) - c_o\hat{z}(t)] \\ \dot{\hat{\theta}}(t) &= \Gamma\xi(t)[y(t) - c_o\hat{z}(t)] \end{aligned}$$

where $k \in \mathbb{R}^n$ and $\Gamma \in \mathbb{R}^{p\times p}$ contain the gain parameters of the algorithm.

As shown in [3], in order to ensure the convergence of this adaptive observer, the gain vector k should be chosen such that the transfer function $c_o(sI - A_o + kc_o)^{-1}\gamma$ is *strictly positive real*, the matrix Γ should be positive definite, and the signals in $\xi(t)$ (and implicitly those in $\Psi(t)$) should satisfy some persistent excitation condition. See [3] for more details.

Because the convergence analysis is based on the strictly positive realness of a *transfer function*, this method is not suitable for the general case of time varying $A(t), C(t)$ matrices.

B. Algorithm 2

Now let us examine the algorithm presented in [4], chapter 5. Only single output systems with constant matrices A, C are considered here. It is assumed that the pair (A, C) is observable. After the static transformation based on coordinate change and output injection, the system corresponding to (2) has the particular form

$$\dot{x}(t) = A_o x(t) + \varphi(t) + \Psi(t)\theta \tag{7a}$$

$$y(t) = c_o x(t) \tag{7b}$$

with

$$A_o = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \; c_o = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$$

Remark that in the original reference [4] the entries in A_o and c_o are arranged in a different way: A_o is transposed with respect to the above definition and the value 1 is at the last position in c_o. Let us use the above defined form in order to be more homogeneous with the other revisited algorithms.

A dynamic transformation of this system has also been introduced in [4]. Its reformulation with the general formula (3) corresponds to the particular choices:

$$F = \begin{bmatrix} -\gamma_2 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\gamma_{n-1} & 0 & \cdots & 1 \\ -\gamma_n & 0 & \cdots & 0 \end{bmatrix}_{(n-1)\times(n-1)}$$

$$G = \begin{bmatrix} -\gamma_2 & 1 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ -\gamma_{n-1} & 0 & \cdots & \ddots & 0 \\ -\gamma_n & 0 & \cdots & \cdots & 1 \end{bmatrix}_{(n-1)\times n}$$

$$H = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ 1 & \cdots & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & \ddots & 0 \\ 0 & \cdots & \cdots & 1 \end{bmatrix}_{n\times(n-1)}$$

where $\gamma_2, \ldots, \gamma_n$ are some chosen values (the requirement on these values will be indicated later), the sizes of the matrices F, G, H are as indicated. Accordingly, the sizes of $\Xi(t)$ and $\Omega(t)$ are respectively $(n-1) \times p$ and $n \times p$.

With this particular transformation applied to (7) and after some straightforward computation, the transformed system (4) becomes

$$\begin{aligned} &\dot{z}(t) = A_o z(t) + \varphi(t) + \gamma\xi^T(t)\theta \\ &y(t) = c_o z(t) \end{aligned}$$

with

$$\gamma = \begin{bmatrix} 1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{bmatrix} \qquad \xi^T(t) = \Xi_1(t) + \Psi_1(t)$$

where $\gamma_2, \ldots, \gamma_n$ are values as used in F, G, the vector Ξ_1 is the first row of Ξ, and Ψ_1 is the first row of Ψ.

Now the presence of the unknown parameter vector θ is also simplified, since it affects the state equation through the scalar quantity $\xi^T(t)\theta$ only. The adaptive observer proposed in [4] for this transformed system takes the form

$$\begin{aligned} &\dot{\hat{z}}(t) = A_o\hat{z}(t) + \varphi(t) + \gamma\xi^T(t)\hat{\theta}(t) + k[y(t) - c_o\hat{z}(t)] \\ &\dot{\hat{\theta}}(t) = \Gamma\xi(t)[y(t) - c_o\hat{z}(t)] \end{aligned}$$

with some gain vector $k \in \mathbb{R}^n$ and a positive definite gain matrix $\Gamma \in \mathbb{R}^{p\times p}$. In order to ensure the convergence of the adaptive observer, the vectors k and γ must be chosen such that the transfer function $c_o(sI - A_o + kc_o)^{-1}\gamma$ is *strictly positive real*. See [4] for more details.

Because the convergence analysis is based on the strictly positive realness of a transfer function, this method is limited to the case of constant A, C matrices.

Some variants of this adaptive observer have been presented in [13], [14].

C. Algorithm 3

In [7], a canonical form for a class of nonlinear systems is proposed for the purpose of adaptive observer design. As an example, the design of adaptive observer for linear time varying (or state-affine) systems in the form of (1) has been presented.

Assume that system (1) is transformed by a coordinate change with output injection into the form

$$\dot{x}(t) = A_o(t)x(t) + \varphi(t) + \Psi(t)\theta \tag{8a}$$
$$y(t) = C_o x(t) \tag{8b}$$

with

$$A_o(t) = \begin{bmatrix} A_{11}(t) & A_{12}(t) \\ A_{21}(t) & A_{22}(t) \end{bmatrix} \qquad C_o = \begin{bmatrix} I_{m\times m} & 0_{m\times\bar{m}} \end{bmatrix}$$

where $\bar{m} = n - m$, $A_{11}(t)$ is a $m \times m$ matrix, $A_{22}(t)$ is a $\bar{m} \times \bar{m}$ *asymptotically stable* matrix, $I_{m\times m}$ is the $m \times m$ identity matrix, $0_{m\times\bar{m}}$ is the $m \times \bar{m}$ zero matrix.

A dynamic transformation has also been proposed in [7]. Its reformulation with the general formula (3) can be made with the choices

$$F(t) = A_{22}(t), \quad G = \begin{bmatrix} 0_{\bar{m}\times m} & I_{\bar{m}\times\bar{m}} \end{bmatrix}, \quad H = \begin{bmatrix} 0_{m\times\bar{m}} \\ I_{\bar{m}\times\bar{m}} \end{bmatrix}$$

where the sizes of $F(t), G$ and H are respectively $\bar{m} \times \bar{m}$, $\bar{m} \times n$ and $n \times \bar{m}$. Accordingly, the sizes of $\Xi(t)$ and $\Omega(t)$ are respectively $\bar{m} \times p$ and $n \times p$.

After this transformation applied to (8), the transformed system (4) becomes

$$\begin{aligned} &\dot{z}(t) = A_o(t)z(t) + \varphi(t) + \begin{bmatrix} I_{m\times m} \\ 0_{\bar{m}\times m} \end{bmatrix} [A_{12}(t)\Xi(t) + \Psi_1(t)]\theta \\ &y(t) = C_o z(t) \end{aligned}$$

where Ψ_1 is the first m rows of Ψ.

If we accordingly divide z and φ into the first m and the last $\bar{m}$ rows,

$$z(t) = \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} \qquad \varphi(t) = \begin{bmatrix} \varphi_1(t) \\ \varphi_2(t) \end{bmatrix}$$

then the transformed system writes

$$\begin{aligned} \dot{z}_1(t) &= A_{11}(t)z_1(t) + A_{12}(t)z_2(t) + \varphi_1(t) \\ &\quad + [A_{12}(t)\Xi(t) + \Psi_1(t)]\theta \end{aligned} \tag{9a}$$
$$\dot{z}_2(t) = A_{21}(t)z_1(t) + A_{22}(t)z_2(t) + \varphi_2(t) \tag{9b}$$
$$y(t) = z_1(t) \tag{9c}$$

This transformed system has the particularity that the unknown parameter θ affects only the equation of $z_1(t)$ which is directly observed through the output $y(t)$. It is in some sense a generalization of the systems in the form of (6).

The adaptive observer proposed in [7] takes the form

$$\begin{aligned} \dot{\hat{z}}_1(t) &= A_{11}(t)\hat{z}_1(t) + A_{12}(t)\hat{z}_2(t) + \varphi_1(t) \\ &\quad + [A_{12}(t)\Xi(t) + \Psi_1(t)]\hat{\theta}(t) + K_1(t)[y(t) - \hat{z}_1(t)] \\ \dot{\hat{z}}_2(t) &= A_{21}(t)y(t) + A_{22}(t)\hat{z}_2(t) + \varphi_2(t) \\ \dot{\hat{\theta}}(t) &= \Gamma[A_{12}(t)\Xi(t) + \Psi_1(t)]^T[y(t) - \hat{z}_1(t)] \end{aligned}$$

with some gain matrices $K_1(t) \in \mathbb{R}^{m\times m}$ and $\Gamma \in \mathbb{R}^{p\times p}$. A Lyapunov function has been designed in [7] for the convergence analysis of this adaptive observer.

This algorithm is applicable to general time varying multi-input multi-output systems. Remark that, due to the time varying nature of the system, in general the coordinate change required to transform system (1) into the form of (8) is also time varying. Such a coordinate change has been derived in [7] . It is based on the solution of a dynamic Riccati equation and requires a time-dependent matrix inversion.

D. Algorithm 4

An adaptive observer has been proposed in [12] directly for systems in the form of (1) without resort to any transformation. Now this algorithm can also be reformulated with transformations as for the previous algorithms, though no such transformations have been used in [12].

Assume that the matrix pair $(A(t), C(t))$ is uniformly observable. Let $K(t) \in \mathbb{R}^{n\times m}$ be a time varying matrix such that

$$A_o(t) = A(t) - K(t)C(t)$$

is exponentially stable. Typically such a $K(t)$ is computed with the aid of a dynamic Riccati equation like in the Kalman filter.

Through output injection, system (1) is transformed to

$$\begin{aligned} \dot{x}(t) &= A_o(t)x(t) + B(t)u(t) + K(t)y(t) + \Psi(t)\theta \\ y(t) &= C(t)x(t) \end{aligned}$$

Remark that *no* coordinate change is required here to obtain this transformed system corresponding to (2). In orther words, the transformation matrix $T = I_{n\times n}$. This simplicity is especially important for time varying systems, since time varying coordinate change is generally not trivial.

Now apply to this system the general dynamic transformation (3) with the very simple choices

$$\begin{aligned} F(t) &= A_o(t) = A(t) - K(t)C(t) \\ G &= I_{n\times n} \\ H &= I_{n\times n} \end{aligned}$$

Then the transformed system takes the form

$$\dot{z}(t) = [A(t) - K(t)C(t)]z(t) + B(t)u(t) + K(t)y(t) \quad (10a)$$

$$y(t) = C(t)z(t) + C(t)\Omega(t)\theta \quad (10b)$$

where $\Omega(t)$ is as defined in (3b) and can be simply described in this particular case by the equation

$$\dot{\Omega}(t) = [A(t) - K(t)C(t)]\Omega(t) + \Psi(t) \quad (11)$$

Notice that the state equation (10a) is independent of the unknown parameters θ. Moreover, the matrix $A_o(t) = A(t) - K(t)C(t)$ is exponentially stable and the vector $\varphi(t) = B(t)u(t) + K(t)y(t)$ contains known signals only. Therefore, the estimation of $z(t)$ is trivial. It is then not difficult to estimate θ from the output equation.

Now the adaptive observer proposed in [12] can be reformulated as

$$\dot{\hat{z}}(t) = [A(t) - K(t)C(t)]\hat{z}(t) + B(t)u(t) + K(t)y(t) \quad (12a)$$

$$\dot{\hat{\theta}}(t) = \Gamma\Omega^T(t)C^T[y(t) - C(t)\hat{z}(t) - C(t)\Omega(t)\hat{\theta}(t)] \quad (12b)$$

where $\Gamma \in \mathbb{R}^{p\times p}$ is a positive definite matrix.

Since this algorithm is formulated here in a form different from that of [12], let us summarize the main lines of its convergence analysis in this new form.

Define the estimation errors $\tilde{z}(t) = \hat{z}(t) - z(t)$ and $\tilde{\theta}(t) = \hat{\theta}(t) - \theta$, and notice that $\dot{\tilde{\theta}}(t) = \dot{\hat{\theta}}(t)$ for constant θ, then from (10) and (12) it is easy to derive

$$\dot{\tilde{z}}(t) = [A(t) - K(t)C(t)]\tilde{z}(t) \quad (13)$$

$$\begin{aligned} \dot{\tilde{\theta}}(t) = &-\Omega^T(t)C^T(t)C(t)\tilde{z}(t) \\ &- \Gamma\Omega^T(t)C^T(t)C(t)\Omega(t)\tilde{\theta}(t) \end{aligned} \quad (14)$$

The error $\tilde{z}(t)$ tends to zero due to the exponential stability of $A_o(t) = A(t) - K(t)C(t)$. In (14) the first term at the right hand side is then vanishing. When this term is omitted, the remaining homogeneous equation is exponentially stable if

$$\int_t^{t+s} \Omega^T(\tau)C^T(\tau)C(\tau)\Omega(\tau)d\tau \geq \alpha I_{p\times p} \quad (15)$$

holds for any t and for some positive constants s and α. See [15], page 72, for a proof. It thus follows that the error $\tilde{\theta}(t)$ driven by the vanishing $\tilde{z}(t)$ tends to zero.

The inequality (15) should be understood as a *persistent excitation condition* on $\Omega(t)$. Since $\Omega(t)$ is generated from $\Psi(t)$ through (11), it is also a condition on $\Psi(t)$ or on $\Phi(t)$ appearing in the original system (1).

The convergence analysis made in [12] is slightly different from the above sketch, since it is based on a different formulation of the same adaptive observer. Some results about the robustness of this algorithm to noises have also been presented in [12].

IV. Conclusion

Though the various revisited algorithms are apparently quite different as presented in their original references, they have been reformulated in this paper in a unified framework. The common steps of static and dynamic transformations allow to achieve some special form simplifying the estimation problem. For the first 3 algorithms, the unknown parameter vector is kept in the transformed state equation, whereas for the last one it is moved to the transformed output equation. This unusual form used in the last algorithm makes the estimation problem extremely simple: the estimation of the transformed state is trivial, and consequently the estimation of the parameter vector is also greatly simplified. It also has a remarkable advantage for time varying systems: no coordinate change is required. The simplicity of this algorithm is especially important when it is used to develop algorithms for nonlinear systems which cannot be linearized. See [11], [8] for such examples.

The revisited algorithms essentially differ by their choices in the static and dynamic transformations. Clearly, it is possible to make other choices in these transformations to design new algorithm variants for particular purposes.

Appendix: comparison between Kalman filter and adaptive observer

Let us consider system (2) in this discussion. The comparison will be made between the adaptive observer algorithm 4 and the Kalman filter applied to the extended system

$$\begin{bmatrix} \dot{x}(t) \\ \dot{\theta}(t) \end{bmatrix} = \begin{bmatrix} A_o(t) & \Psi(t) \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \theta(t) \end{bmatrix} + \begin{bmatrix} \varphi(t) \\ 0 \end{bmatrix} \quad (16a)$$

$$y(t) = \begin{bmatrix} C_o(t) & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \theta(t) \end{bmatrix} \quad (16b)$$

Indeed this is still a linear system, thus the Kalman filter is applicable for joint estimation of $x(t)$ and θ. However, the convergence of this Kalman filter requires the uniform observability [16] of the *extended* system. Remark that even in the case of constant matrices A_o, C_o, the extended system is *time varying* due to the time varying $\Psi(t)$. Its uniform observability is defined with the aid of the Gramian observability matrix as derived in the following.

Let $\Phi(t, t_0) \in \mathbb{R}^{(n+p)\times(n+p)}$ be the transition matrix of the extended system (16) (please do not confuse this notation with the Φ used in (1)). Then,

$$\frac{d}{dt}\Phi(t, t_0) = \begin{bmatrix} A_o(t) & \Psi(t) \\ 0_{p\times n} & 0_{p\times p} \end{bmatrix} \Phi(t, t_0)$$
$$\Phi(t_0, t_0) = I_{(n+p)\times(n+p)}$$

It is easy to check that

$$\Phi(t, t_0) = \begin{bmatrix} \Phi_{11}(t, t_0) & \Phi_{12}(t, t_0) \\ 0_{p\times n} & I_{p\times p} \end{bmatrix}$$

with

$$\frac{d}{dt}\Phi_{11}(t, t_0) = A_o(t)\Phi_{11}(t, t_0)$$
$$\Phi_{11}(t_0, t_0) = I_{n\times n}$$
$$\frac{d}{dt}\Phi_{12}(t, t_0) = A_o(t)\Phi_{12}(t, t_0) + \Psi(t)$$
$$\Phi_{12}(t_0, t_0) = 0_{n\times p}$$

It is thus clear that $\Phi_{11}(t, t_0)$ is the transition matrix related to the matrix $A_o(t)$. The matrix $\Phi_{12}(t, t_0)$ coincides with $\Omega(t)$ as generated with (11) (remind that $A_o = A(t) - K(t)C(t)$ for this adaptive observer).

The uniform observability of the extended system (16) is conditioned by the uniform positive definiteness of the $(n + p) \times (n + p)$ Gramian matrix

$$W(t, s) = \int_t^{t+s} \Phi^T(\tau, t)[C_o(\tau)\ 0]^T[C_o(\tau)\ 0]\Phi(\tau, t)d\tau$$
$$= \begin{bmatrix} W_{11}(t, s) & W_{12}(t, s) \\ W_{12}^T(t, s) & W_{22}(t, s) \end{bmatrix}$$

with

$$W_{11}(t, s) = \int_t^{t+s} \Phi_{11}^T(\tau, t)C_o^T(\tau)C_o(\tau)\Phi_{11}(\tau, t)d\tau$$
$$W_{12}(t, s) = \int_t^{t+s} \Phi_{11}^T(\tau, t)C_o^T(\tau)C_o(\tau)\Phi_{12}(\tau, t)d\tau$$
$$W_{22}(t, s) = \int_t^{t+s} \Phi_{12}^T(\tau, t)C_o^T(\tau)C_o(\tau)\Phi_{12}(\tau, t)d\tau$$

The persistent excitation (15) coincides with the uniform positive definiteness of $W_{22}(t, s)$, whereas the uniform observability of the matrix pair $(A_o(t), C_o(t))$ coincides with that of $W_{11}(t, s)$.

Though the positive definiteness of $W_{11}(t, s)$ and $W_{22}(t, s)$ imply the positive definiteness of $W(t, s)$, the *uniform* positive definiteness of $W(t, s)$ cannot be implied by those of the sub-matrices. Therefore, the *uniform* positive definiteness of the $(n + p) \times (n + p)$ Gramian matrix $W(t, s)$ should be checked to ensure the convergence of the Kalman filter, whereas for the adaptive observer only the sub-matrices $W_{11}(t, s)$ and $W_{22}(t, s)$ need to be checked.

For numerical implementation, the Kalman filter of the extended system (16) involves a $(n + p) \times (n + p)$ Riccati equation, whereas usually the matrix $K(t)$ stabilizing $A_o(t) = A(t)-K(t)C(t)$ required by the adaptive observer is based on a $n \times n$ Riccati equation.

REFERENCES

[1] G. Lüders and K. S. Narendra, "An adaptive observer and identifier for a linear system," *IEEE Trans. on Automatic Control*, vol. 18, pp. 496–499, oct 1973.

[2] G. Kreisselmeier, "Adaptive observers with exponential rate of convergence," *IEEE Trans. on Automatic Control*, vol. 22, no. 1, pp. 2–8, feb 1977.

[3] G. Bastin and M. Gevers, "Stable adaptive observers for nonlinear time varying systems," *IEEE Trans. on Automatic Control*, vol. 33, no. 7, pp. 650–658, 1988.

[4] R. Marino and P. Tomei, *Nonlinear control design*, ser. Information and system sciences. London, New York: Prentice Hall, 1995.

[5] R. Rajamani and K. Hedrick, "Adaptive observer for active automotive suspensions - theory and experiment," *IEEE Trans. on Control Systems Technology*, vol. 3, no. 1, pp. 86–93, 1995.

[6] Y. M. Cho and R. Rajamani, "A systematic approach to adaptive observer synthesis for nonlinear systems," *IEEE Trans. on Automatic Control*, vol. 42, no. 4, pp. 534–537, apr 1997.

[7] G. Besançon, "Remarks on nonlinear adaptive observer design," *Systems and Control Letters*, vol. 41, no. 4, pp. 271–280, 2000.

[8] A. Xu and Q. Zhang, "Nonlinear system fault diagnosis based on adaptive estimation," *Automatica*, vol. 40, no. 7, pp. 1181–1193, 2004.

[9] A. S. Poznyak and J. C. Martinez, "Variable structure robust state and parameter estimator," *International Journal of Adaptive Control and Signal Processing*, vol. 15, no. 4, pp. 179–208, 2001.

[10] J. C. Martinez and A. S. Poznyak, "Switching structure state and parameter estimator for MIMO non-linear robust control," *International Journal of Control*, vol. 74, no. 2, pp. 175–189, 2001.

[11] Q. Zhang, A. Xu, and G. Besançon, "An efficient nonlinear adaptive observer with global convergence," in *13th IFAC/IFORS Symposium on System Identification (SYSID)*, Rotterdam, 2003, pp. 1737–1742.

[12] Q. Zhang, "Adaptive observer for multiple-input-multiple-output (MIMO) linear time varying systems," *IEEE Trans. on Automatic Control*, vol. 47, no. 3, pp. 525–529, 2002.

[13] R. Marino and P. Tomei, "Adaptive observers with arbitrary exponential rate of convergence for nonlinear systems," *IEEE Trans. on Automatic Control*, vol. 40, no. 7, pp. 1300–1304, 1995.

[14] R. Marino and G. L. Santosuosso, "Robust adaptive observers for nonlinear systems with bounded disturbances," in *38th IEEE Conference on Decision and Control (CDC'99)*, Phoenix, Arizona, 1999.

[15] K. S. Narendra and A. M. Annaswamy, *Stable adaptive systems*. Boston: Prentice Hall, 1989.

[16] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, ser. Mathematics in Science and Engineering. New York: Academic Press, 1970, vol. 64.