2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Generality of Functional Observer Structures

T. L. Fernando, L. S. Jennings, H. M. Trinh

***Abstract*— Functional observers estimate a linear function of the state vector directly without having to estimate all the individual states. In the past various observer structures have been employed to design such functional estimates. In this paper we discuss the generality of those various observer structures and prove the conditions under which those observer structures are unified. The paper also highlights and clarifies the need to remove the self-convergent states from the system and also from the functions to be estimated before proceeding with the design of a functional observer or else incorrect conclusions regarding the existence of functional observers can be arrived at.**

I. Introduction

Functional observers can play an important role in state feedback controller implementation because a control law can be directly estimated from input and output data using a functional observer. The main advantage of employing a functional observer lies in its ability to directly estimate a given linear function of the state vector without having to estimate all the individual states whereas a state observer design scheme cannot. The direct estimation of a linear function allows the observer structure to have minimal dimension, always an order less or equal to the reduced-order Luenberger observer. Designing the simplest possible order observer to estimate a given linear function from a practical point of view allows its implementation with the least possible components bringing cost benefits while ensuring a circuit implementing the observer dynamics is the simplest possible. This salient feature of possible reduced-order for a functional observer has motivated researchers around the world to find ways to systematically design minimum order functional observers since the concept was first introduced by Luenberger [1] in 1966.

Existence conditions for functional observers were known as early as 1966, see [1]- [4], those reported conditions are in terms of the observer parameters. In this paper we will show that those reported conditions in early work are necessary and sufficient only under certain assumptions, in particular in the absence of self-convergent states. The most complete algebraic conditions for the existence of a predetermined order functional observer was reported by Darouach in [12], those reported conditions are in terms of the system matrix A, output matrix C and also in terms of a matrix representing the linear functions to be estimated L_0, the predetermined order of the observer being the number of rows of L_0. Moreno in [13] reported functional observer existence conditions with no reference to the order of the observer, the condition reported in [13] is also useful in the sense it can establish the existence of a functional observer in terms of known parameters A, C and L_0. The observer structures employed by Darouach in [12] and Moreno in [13] are however different. Tsui in [8] has also reported some early studies on the reduction of the order of functional observers, again the structure of the observer employed in [8] is different to the two different structures employed by Darouach and Moreno. In the literature there are at least three different structures that have been employed in the studies of functional observers. In this paper we prove under stated assumptions which can be made without loss of generality that the three types of structures for functional state estimation share the same level of generality.

In summary, the contributions of this paper are:

I Unify various functional observer structures reported in the literature under the stated assumptions in the paper.

II Highlight and clarify the role of self-convergent states in the existence of functional observers and relaxes the assumption on Controllability of the system in designing an observer of the most general structure.

II. Problem Statement

Consider a linear time-invariant system described by

$$\dot{x}(t) = Ax(t) + Bu(t), \tag{1a}$$

$$y(t) = Cx(t), \tag{1b}$$

$$z_0(t) = L_0 x(t), \tag{1c}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^p$ are the state, input and the output vectors, respectively and $z_0(t) \in \mathbb{R}^r$ is the vector to be estimated. $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$ and $L_0 \in \mathbb{R}^{r\times n}$ are known constant matrices. We assume

$$(\mathbf{A1}): \quad \text{The triple } (A, C, L_0) \text{ is Functional Observable see [18], but the pair } (A, C) \text{ is not necessarily Observable.} \tag{2}$$

We also make the following assumptions and argue later no loss of generality.

$$(\mathbf{A2}): \quad \operatorname{rank}(C) = p. \tag{3}$$

$$(\mathbf{A3}): \quad \operatorname{rank}(L_0) = r. \tag{4}$$

$$(\mathbf{A4}): \quad \operatorname{rank}\left(\begin{bmatrix} C \\ L_0 \end{bmatrix}\right) = p + r. \tag{5}$$

T. L. Fernando is with the School of Electrical, Electronic and Computer Engineering, University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia. E-mail: tyrone@ee.uwa.edu.au

L. S. Jennings is with the School of Mathematics and Statistics, University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia. E-mail: les@maths.uwa.edu.au

H. M. Trinh is with the School of Engineering, Deakin University, Geelong, VIC 3217, Australia. E-mail: hmt@deakin.edu.au

978-1-61284-799-3/11/$26.00 ©2011 IEEE

$$(\mathbf{A5}): \text{ The system has no stable uncontrollable states.} \quad (6)$$

Let $z(t) \in \mathbb{R}^q$

$$z(t) = Lx(t), \quad (7)$$

where $L \in \mathbb{R}^{q \times n}, q \geq r$ is a built matrix such that the first r rows of L are the rows of L_0. We make further assumptions,

$$(\mathbf{A6}): \quad \text{rank}\,(L) = q. \quad (8)$$

$$(\mathbf{A7}): \quad \text{rank}\left(\begin{bmatrix} C \\ L \end{bmatrix}\right) = p + q, \text{ (hence } q \leq n - p). \quad (9)$$

A functional observer of order q has the following most general form

$$\dot{w}(t) = Nw(t) + Jy(t) + Hu(t), \quad (10a)$$

$$\hat{z}(t) = Qw(t) + Ey(t), \quad (10b)$$

$$\hat{z}_0(t) = \left[\begin{array}{c|c} \mathbf{I}_r & \mathbf{0}_{r\times(q-r)} \end{array}\right] \hat{z}(t) \text{ or } L_0 = \left[\begin{array}{c|c} \mathbf{I}_r & \mathbf{0}_{r\times(q-r)} \end{array}\right] L, \quad (10c)$$

where $w(t) \in \mathbb{R}^q$. In this paper we will show that there is no loss of generality in the observer structure (10a)-(10c) in assigning: (I) $Q = \mathbf{I}_q$ with no assumed structure for any of the other observer parameters N, J, H and E or (II) N diagonal and with no assumed structure for any of the other observer parameters J, H, Q and E.

III. Main Result

In the literature there has been at least three types of observer structures proposed to design functional observers:

I. The form (10a)-(10c) with no assumed structure for any of the observer parameters N, J, H, Q and E, see [1]- [6], [10] and [15]- [17].
II. The form (10a)-(10c) with no assumed structure for the observer parameters N, J, H and E, and assigning $Q = \mathbf{I}_q$, see [12] and [18].
III. The form (10a)-(10c) with no assumed structure for the observer parameters J, H, Q and E, and assuming a diagonal structure for N, see [7]- [9], [11] and [14].

According to the problem formulation in this paper, in particular (A1)-(A7), all three observer structures are equally general. First we state the rationale for the assumptions (A1)-(A7). Assumption (A1) implies that asymptotic estimation of $z_0(t)$ is possible (see [13] and [18]), i.e., if (A1) is not satisfied then $z_0(t)$ cannot be estimated and an observer of the form (10a)-(10c) does not exist. The assumption (A2) implies that all p outputs are linearly independent (linearly dependent outputs can be ignored as those outputs provide no extra information), and (A3) implies all r functions being estimated are linearly independent and any linearly dependent functions are not considered because those can be obtained from the independent estimates, and furthermore (A4) implies that outputs and the functions being estimated are linearly independent because any functions which are linearly dependent on the output can be obtained from linearly combining the outputs without having to estimate using a dynamical system. Same can be said about assumptions (A6) and (A7) as per (A3) and (A4) because $Lx(t)$ also represents functions that can be estimated. The assumption (A5) is about uncontrollable self-convergent states, the asymptotic value of uncontrollable but self-convergent states are zero, those states play no part in an asymptotic estimation of $z_0(t)$, and if present, those states can be removed from the system and the function $z_0(t)$. Furthermore, given (A2)-(A5), clearly a static observer can play no part in estimating $z_0(t)$ (i.e., $z_0(t)$ cannot be obtained by linearly combining the outputs) hence the lower bound for the order of a functional observer to asymptotically estimate $z_0(t)$ is r. Since the lower bound for a functional observer is r, it follows from (A1) that the function $z_0(t)$ can be estimated asymptotically with an observer of order q.

An intermediate error $\epsilon(t) \in \mathbb{R}^q$ and its dynamics can be written as follows:

$$\epsilon(t) = w(t) - Px(t), \quad (11)$$

$$\dot{\epsilon}(t) = N\epsilon(t) + (NP - PA + JC)\,x(t) + (H - PB)\,u(t), \quad (12)$$

where $P \in \mathbb{R}^{q \times n}$. The estimation errors $e(t) \in \mathbb{R}^q$ and $e_0(t) \in \mathbb{R}^r$ can be written as

$$e(t) = \hat{z}(t) - z(t) = Q\epsilon(t) + (QP + EC - L)x(t), \quad (13)$$

$$e_0(t) = \hat{z}_0(t) - z_0(t) = \left[\begin{array}{c|c} \mathbf{I}_r & \mathbf{0}_{r\times(q-r)} \end{array}\right] e(t). \quad (14)$$

Lemma 1: If

$$QP + EC - L = \mathbf{0}_{q\times n} \quad (15)$$

then $\text{rank}(Q) = \text{rank}(P) = q$ and hence Q is invertible.

Proof: From (15)

$$QP = \left[\begin{array}{c|c} \mathbf{I}_q & -E \end{array}\right] \begin{bmatrix} L \\ C \end{bmatrix}. \quad (16)$$

From (A2), (A3) and (A4) the matrix $\begin{bmatrix} L\,;\,C \end{bmatrix}$ is full row rank, and also from the fact that $\text{rank}(M_1 M_2) = \text{rank}(M_1)$ if M_2 is full row rank, the following can be written

$$\text{rank}\left(\left[\begin{array}{c|c} \mathbf{I}_q & -E \end{array}\right] \begin{bmatrix} L \\ C \end{bmatrix}\right) = \text{rank}\left(\left[\begin{array}{c|c} \mathbf{I}_q & -E \end{array}\right]\right) = q. \quad (17)$$

From (16) and (17)

$$\text{rank}(QP) = q. \quad (18)$$

From $\text{rank}(M_1 M_2) \leq \min\Big(\text{rank}(M_1), \text{rank}(M_2)\Big)$ and from (18)

$$\text{rank}(Q) \geq q \quad \text{and} \quad \text{rank}(P) \geq q, \quad (19)$$

and as both have q rows

$$\text{rank}(Q) = q = \text{rank}(P). \quad \triangle\triangle\triangle \quad (20)$$

Theorem 1: $e(t) \to \mathbf{0}_{q\times 1}$ as $t \to \infty$ for any $x(0), w(0)$ and $u(t)$ and hence $e_0(t) \to \mathbf{0}_{r\times 1}$ as $t \to \infty$ for any $x(0), w(0)$ and $u(t)$ iff

$$\epsilon(t) \to \mathbf{0}_{q\times 1} \text{ as } t \to \infty, \quad (21a)$$

$$QP + EC - L = \mathbf{0}_{q\times 1}, \quad (21b)$$

Proof: If (21a)-(21b) are satisfied then $e(t) \to \mathbf{0}_{q\times 1}$ as $t \to \infty$ for any $u(t)$ and $\epsilon(0)$ (hence for any $x(0)$ and $w(0)$ as well). If (21b) is not satisfied then the uncontrollable states of $x(t)$ (note they are all unstable) approach ∞ and the controllable states of $x(t)$ can be arbitrarily chosen using $u(t)$ to make $e(t) \not\to \mathbf{0}_{q\times 1}$. If (21a) is not satisfied then $e(t) \not\to \mathbf{0}_{q\times 1}$ because Q is full rank. △△△

Theorem 2: $\epsilon(t) \to \mathbf{0}_{q\times 1}$ as $t \to \infty$ for any $x(0), w(0)$ and $u(t)$ iff

$$N \text{ is Hurwitz,}$$
$$PA - NP - JC = \mathbf{0}_{q\times 1}, \tag{22a}$$
$$H - PB = \mathbf{0}_{q\times m}. \tag{22b}$$

Proof: If (22a)-(22b) are satisfied and also if N is Hurwitz then $\epsilon(t) \to \mathbf{0}_{q\times 1}$ for any $u(t)$ and $\epsilon(0)$ (hence for any $x(0)$ and $w(0)$ as well). If N has some of its eigenvalues $\lambda_i \in \mathbb{C}$ such that $\Re(\lambda_i) \geq 0$ then even for $x(0) = \mathbf{0}_{n\times 1}$ and $u(t) = \mathbf{0}_{m\times 1}$, $\epsilon(t) \to \infty$ as $t \to \infty$. If (22b) is not satisfied then there exists a $u(t)$ that makes $\epsilon(t) \not\to \mathbf{0}_{q\times 1}$ as $t \to \infty$. If (22a) is not satisfied then the uncontrollable states of $x(t)$ (note they are all unstable) approach ∞ and the controllable states of $x(t)$ can be arbitrarily chosen using $u(t)$ to make $\epsilon(t) \not\to \mathbf{0}_{q\times 1}$. △△△

Remark 1: A theorem similar to Theorems 1 and 2 in this paper appeared in [18], however the theorem in [18] is based on an observer structure (10a)-(10c) with $Q = \mathbf{I}_q$, whereas Theorems 1 and 2 of this paper are applicable for the most general observer structure (10a)-(10c) with Q not necessarily chosen as $Q = \mathbf{I}_q$.

Remark 2: A proof similar to that of Theorem 2 is reported in [2] and [3], however it is based on a Controllability assumption of the pair (A, B). While a proof based on a Controllability assumption is valid, such an assumption on Controllability is unnatural for an observer design problem, and it is clear from the proof of Theorem 2 such an assumption on Controllability is not required.

Remark 3: Note the order in which the necessity of the proof of Theorem 2 is established, first the necessity of N being Hurwitz is established followed by the necessity of (22b). Satisfaction of (22b) makes all $u(t) \in \mathbb{R}^m$ belong to the nullspace of $H - TB$. Finally the necessity of (22a) is established, since $u(t)$ can take any value in $\mathbb{R}^m$ the controllable states of $x(t)$ can be chosen arbitrarily, and furthermore the unstable and uncontrollable states can be made to take any value at any given time instant by choosing its initial condition.

Corollary 1: $e(t) \to \mathbf{0}_{q\times 1}$ as $t \to \infty$ for any $x(0), w(0)$ and $u(t)$ and hence $e_0(t) \to \mathbf{0}_{r\times 1}$ as $t \to \infty$ for any $x(0), w(0)$ and $u(t)$ iff conditions in Theorems 1 and 2 are satisfied.

Theorem 3: $e(t) \to \mathbf{0}_{q\times 1}$ with an arbitrary rate of convergence as $t \to \infty$ for any $x(0), w(0)$ and $u(t)$, and hence $e_0(t) \to \mathbf{0}_{r\times 1}$ with an arbitrary rate of convergence as $t \to \infty$ for any $x(0), w(0)$ and $u(t)$ iff the following two conditions reported in Lemmas 1 and 2 of [12] are satisfied:

$$\text{rank}\begin{bmatrix} LA \\ CA \\ C \\ L \end{bmatrix} = \text{rank}\begin{bmatrix} CA \\ C \\ L \end{bmatrix}, \tag{23a}$$
$$\text{rank}\begin{bmatrix} sL - LA \\ CA \\ C \end{bmatrix} = \text{rank}\begin{bmatrix} CA \\ C \\ L \end{bmatrix}, s \in \mathbb{C}. \tag{23b}$$

Proof: Since Q is invertible (see Lemma 1 of this paper), from (21b) we have

$$P = Q^{-1}L - Q^{-1}EC, \tag{24}$$

Post multiplying (22a) by a full row rank matrix $\begin{bmatrix} L^+ & \mathbf{I}_n - L^+L \end{bmatrix}$ the following two equations can be written:

$$(PA - NP - JC)\,L^+ = \mathbf{0}_{q\times q}, \tag{25a}$$
$$(PA - NP - JC)\,(\mathbf{I}_n - L^+L) = \mathbf{0}_{q\times n}. \tag{25b}$$

Using (24), (25a) can be written as

$$QNQ^{-1} = LAL^+ - \begin{bmatrix} E & K \end{bmatrix}\begin{bmatrix} CAL^+ \\ CL^+ \end{bmatrix} \tag{26}$$

where

$$K = QJ - QNQ^{-1}E. \tag{27}$$

Now consider (25b), it can be rewritten as:

$$L\overline{A} = \begin{bmatrix} E & K \end{bmatrix}\Sigma \tag{28}$$

where $\overline{A} = A(\mathbf{I}_n - L^+L)$, $\overline{C} = C(\mathbf{I}_n - L^+L)$ and $\Sigma = \begin{bmatrix} C\overline{A} \\ \overline{C} \end{bmatrix}$. Now from proof of Lemma 1 in [12], it directly follows that E and K in (28) have a solution iff (23a) is satisfied, and from proof of Lemma 2 in [12], it directly follows that QNQ^{-1} in (26) can be made Hurwitz (N can be made Hurwitz as well since $\text{eig}(N) = \text{eig}(QNQ^{-1})$) iff (23b) is satisfied. △△△

Remark 4: Notable differences of Theorem 3 to that in [12] is that Q not being chosen an Identity matrix and also makes no assumption on the structure of Q at the outset, and also L not being an original set of rows, here L is a built matrix which includes L_0 in its rows whereas in [12] L is given at $L = L_0$. Furthermore, the order of the functional observer considered in this paper is q, $r \leq q \leq (n-p)$ and the assumption (A1) implies that Conditions (23a) and (23b) can always be satisfied whereas the order of the functional observer considered in [12] is r and the assumption (A1) does not necessarily imply that Conditions (23a) and (23b) are satisfied. How to find the mnimum value of q and also how to find a matrix L with the minimum number of rows is reported in [18].

Remark 5: According Lemma 1 the matrix Q can be any invertible matrix, hence there is no loss of generality in choosing $Q = \mathbf{I}_q$. The method of choosing the observer matrices N, J, H and E is then identical to the procedure described in [12].

Corollary 2: Since $L_0x(t)$ can be estimated asymptotically with a q order observer, $SL_0x(t)$ where S is an invertible matrix can also be estimated asymptotically with a q order observer, hence $L_0x(t)$ and $SL_0x(t)$ can be regarded as equivalent functions. The result can be easily proven by replacing L_0 with SL_0 in the conditions of Theorem 3 or by merely taking a linear combination of the estimates of $L_0x(t)$ to obtain $SL_0x(t)$.

Now consider a functional observer with a diagonal structure (Jordan form) for $\tilde{N}$:

$$\dot{\tilde{w}}(t) = \tilde{N}\tilde{w}(t) + \tilde{J}y(t) + \tilde{H}u(t), \tag{29a}$$
$$\hat{z}(t) = \tilde{Q}\tilde{w}(t) + \tilde{E}y(t), \tag{29b}$$
$$\hat{z}_0(t) = \left[\begin{array}{c|c} \mathbf{I}_r & \mathbf{0}_{r\times(q-r)} \end{array}\right]\hat{z}(t), \tag{29c}$$

where $\tilde{w}(t) \in \mathbb{R}^q$, $\tilde{N} \in \mathbb{R}^{q\times q}$, $\tilde{J} \in \mathbb{R}^{q\times p}$, $\tilde{H} \in \mathbb{R}^{q\times m}$, $\tilde{Q} \in \mathbb{R}^{q\times q}$ and $\tilde{E} \in \mathbb{R}^{q\times p}$. The observer structure proposed in [8] to estimate $z(t)$ is of the form (29a)-(29c). However, in [8] the two equations (29b) and (29c) are combined to give just one equation.

Let us also consider a functional observer with $Q = \mathbf{I}_q$:

$$\dot{\bar{w}}(t) = \bar{N}\bar{w}(t) + \bar{J}y(t) + \bar{H}u(t), \tag{30a}$$
$$\hat{z}(t) = \bar{w}(t) + \bar{E}y(t), \tag{30b}$$
$$\hat{z}_0(t) = \left[\begin{array}{c|c} \mathbf{I}_r & \mathbf{0}_{r\times(q-r)} \end{array}\right]\hat{z}(t), \tag{30c}$$

where $\bar{w}(t) \in \mathbb{R}^q$, $\bar{N} \in \mathbb{R}^{q\times q}$, $\bar{J} \in \mathbb{R}^{q\times p}$, $\bar{H} \in \mathbb{R}^{q\times m}$ and $\bar{E} \in \mathbb{R}^{q\times p}$. The observer structure proposed in [12] is of the form (30a)-(30c) with $q = r$. The structure proposed in [18] is of the form (30a)-(30c).

Theorem 4: Under the stated assumptions A1-A7, all three observer structures (10a)-(10c), (29a)-(29c) and (30a)-(30c) are equivalent similarity transformed systems.

Proof: As $\tilde{Q}$ is invertible, choosing $\tilde{Q}\tilde{w} = \bar{w}$, similarity transforms (30a)-(30b) to (29a)-(29b),

$$\dot{\tilde{w}}(t) = \tilde{Q}^{-1}\bar{N}\tilde{Q}\tilde{w}(t) + \tilde{Q}^{-1}\bar{J}y(t) + \tilde{Q}^{-1}\bar{H}u(t), \tag{31a}$$
$$\hat{z}(t) = \tilde{Q}\tilde{w}(t) + \tilde{E}y(t), \tag{31b}$$

so that $\tilde{N} = \tilde{Q}^{-1}\bar{N}\tilde{Q}$, $\tilde{J} = \tilde{Q}^{-1}\bar{J}$, $\tilde{H} = \tilde{Q}^{-1}\bar{H}$, $\tilde{E} = \bar{E}$ where $\tilde{Q}$ is the matrix of eigenvectors of $\bar{N}$. The observer structure (10a)-(10b) can be similarity transformed into either (29a)-(29b) or (30a)-(30b) and vice versa, $\bar{w} = Qw$ transforms (10a)-(10b) to (30a)-(30b) so that $\bar{N} = QNQ^{-1}$, $\bar{J} = QJ$, $\bar{H} = QH$, $\bar{E} = E$. Similarly $\tilde{Q}\tilde{w} = Qw$ or $\tilde{w} = \tilde{Q}^{-1}Qw$ transforms (10a)-(10b) to (29a)-(29b) so that $\tilde{N} = \tilde{Q}^{-1}QNQ^{-1}\tilde{Q}$, $\tilde{J} = \tilde{Q}^{-1}QJ$, $\tilde{H} = \tilde{Q}^{-1}QH$ and $\tilde{E} = E$. $\triangle\triangle\triangle$

Remark 6: While all three observer structures are equivalent similarity transformed systems, the general case (10a)-(10c) is an unnecessary similarity transformation from both of the other structures. The general case (10a)-(10c) does not exploit the inherent redundancies in its structure.

IV. Numerical Examples

In this section we present three examples to highlight the three main contributions of this paper.

Example 1 - *On the uncontrollable self-convergent states and relaxation of the Controllability assumption.*

Consider the following third-order uncontrollable and unobservable system

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ -2 & 2 & -1 \\ 1 & -1 & -3 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} u(t)$$

and

$$y(t) = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}$$

Let the function to be estimated given by

$$z_0(t) = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}$$

Let us first design a functional observer without removing the self-convergent state. Conditions in Theorem 3 are not satisfied hence conditions in Theorems 1 and 2 cannot also be satisfied and one would arrive at the conclusion that a first-order functional observer cannot be designed. Furthermore, let us consider the functional observer existence conditions reported in [13] for assigning arbitrary poles

$$\text{rank} \begin{bmatrix} sI - A \\ C \\ L_0 \end{bmatrix} = \text{rank} \begin{bmatrix} sI - A \\ C \end{bmatrix} \quad \forall \quad s \in \mathbb{C}. \tag{32}$$

The RHS of (32) is 2 when $s = -1$ whereas its LHS is 3, so clearly (32) is not satisfied for all $s \in \mathbb{C}$ and one would arrive at a conclusion that an asymptotic observer with arbitrary poles does not exist to estimate $z_0(t)$. However, if the problem is approached according to the formulation in this paper, in particular by removing the self-convergent state then a first-order observer can be designed as detailed below.

Clearly the state $x_1(t)$ is self-convergent. Based on the stated assumptions in this paper $x_1(t)$ should be removed from the system and the function to be estimated and that will not alter the asymptotic value of $y(t)$, $x_2(t)$, $x_3(t)$ and also $z_0(t)$. The modified system then becomes

$$\begin{bmatrix} \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & -3 \end{bmatrix} \begin{bmatrix} x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(t)$$

and

$$y(t) = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_2(t) \\ x_3(t) \end{bmatrix}$$

The function to be estimated becomes

$$z_0(t) = \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} x_2(t) \\ x_3(t) \end{bmatrix}$$

It is now easy to verify that conditions in Theorem 3 are satisfied and therefore a first-order functional observer can be designed to estimate $z_0(t) = x_2(t) + x_3(t)$ which has the same asymptotic value of $x_1(t) + x_2(t) + x_3(t)$. By considering the eigenvalues of A it is easy to see that (32) is also satisfied implying the existence of a functional observer. This example illustrates the need to remove the self-convergent states from the system and also from the function to be estimated before proceeding with the design of a functional observer or else Theorems 1 and 2 and consequently Theorem 3 as well may arrive at an incorrect conclusion. The same can be said about the existence condition (32). Theorems 1, 2 and 3, and also the existence condition (32) are all valid results but require the removal of the self-convergent states at the outset and does not require a Controllability assumption on the pair (A, B).

Example 2 - *On the existence of all three functional observer structures.*

Consider the numerical example reported in [4] and [5] where A, B, C and L_0 are as follows:

$$A = \begin{bmatrix} 1 & 0 & 0 & -1 & 1 \\ -2 & 0 & 0 & 2 & -1 \\ 0 & 1 & -1 & 0 & -1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 \\ -1 & 2 \\ -1 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix},$$

$C = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$, and

$L_0 = -\frac{1}{4}\begin{bmatrix} 6 & 1 & 1 & 2 & 2 \\ -2 & 3 & -1 & 2 & 2 \end{bmatrix}$.

Following the design algorithm in [4] an observer structure of the form (10a), (10c) below is reported in [4]

$$\dot{w}(t) = \begin{bmatrix} -7.8 & 0.2 \\ -72.2 & -0.2 \end{bmatrix} w(t) + \begin{bmatrix} -57 & -16 \\ -662 & -175 \end{bmatrix} y(t) + \begin{bmatrix} -9.4 & -11.2 \\ -85.6 & -82.8 \end{bmatrix} u(t),$$

$$\hat{z}(t) = \begin{bmatrix} -1.5 & -0.25 \\ 0.5 & -0.75 \end{bmatrix} w(t) + \begin{bmatrix} -31.25 & -6.25 \\ -48.75 & -10.75 \end{bmatrix} y(t).$$

It is easy to verify that the above dynamical system satisfies conditions in Theorem 2, hence the dynamical system is an asymptotic functional observer. Furthermore, both poles of the observer are at $s = -4$. Now let us consider designing a functional observer of the form (29a)-(29c), this can be performed using the method proposed by Darouach [12] because Lemma 1 and Lemma 2 of [12] are satisfied and according to the algorithm in [12] the following functional observer of the form (30a)-(30c) can be derived to place both poles of the observer at $s = -4$.

$$\dot{\bar{w}}(t) = \begin{bmatrix} -17.75 & 6.25 \\ -30.25 & 9.75 \end{bmatrix} \bar{w}(t) + \begin{bmatrix} 251 & 67.75 \\ 468 & 123.25 \end{bmatrix} y(t) + \begin{bmatrix} 35.5 & 37.5 \\ 59.5 & 56.5 \end{bmatrix} u(t),$$

$$\hat{z}(t) = \bar{w}(t) + \begin{bmatrix} -31.25 & -6.25 \\ -48.75 & -10.75 \end{bmatrix} y(t).$$

It is also possible to design a functional observer of the form (29a)-(29c), as follows:

$$\dot{\tilde{w}}(t) = \begin{bmatrix} -4 & 1 \\ 0 & -4 \end{bmatrix} \tilde{w}(t) + \begin{bmatrix} -606.6 & -163.7 \\ 1271.7 & 389.7 \end{bmatrix} y(t) + \begin{bmatrix} -85.7896 & -90.6228 \\ 280.9307 & 392.6988 \end{bmatrix} u(t),$$

$$\hat{z}(t) = \begin{bmatrix} -0.4138 & 0 \\ -0.9104 & -0.0662 \end{bmatrix} \tilde{w}(t) + \begin{bmatrix} -31.25 & -6.25 \\ -48.75 & -10.75 \end{bmatrix} y(t).$$

This example demonstrates the existence of all three observer structures consistent with Theorem 4 of this paper, it is also clear that all three functional observer structures are related through similarity transformations and the variables $w(t)$, $\tilde{w}(t)$ and $\bar{w}(t)$ are related as per Theorem 4:

$$\bar{w} = \begin{bmatrix} -1.5 & -0.25 \\ 0.5 & -0.75 \end{bmatrix} w, \ \bar{w} = \begin{bmatrix} -0.4138 & 0 \\ -0.9104 & -0.0662 \end{bmatrix} \tilde{w}.$$

V. Conclusion

This paper highlights the importance of removing the self-convergent states from the system before proceeding with the design of a functional observer because the presence of self-convergent states can lead to incorrect conclusions regarding the existence of an observer. An assumption on Controllability in designing an observer of the most general form has also been relaxed. The paper also clearly establishes assumptions that can be made without loss of generality which unifies various observer structures that can be used in functional state estimation.

References

[1] D. G. Luenberger, "Observers for multivariable systems," *IEEE Trans. Automat. Contr.*, vol. 11, pp. 190-197, 1966.

[2] D. G. Luenberger, "An Introduction to Observers," *IEEE Trans. Automat. Contr.*, vol. 16, pp. 596-602, 1971.

[3] T. E. Fortmann and D. Williamson, "Design of low-order observers for linear feedback control laws", *IEEE Trans. Automat. Contr.*, vol. 17, pp. 301-308, 1972.

[4] J.R. Roman and T.E. Bullock, "Design of minimal-order stable observers for linear functions of the state via realization theory," *IEEE Trans. Automat. Contr.*, vol. 20, pp. 613-622, 1975.

[5] K. Furuta and S. Kawaji, "Linear functional observer with possible minimal dimension," *IEEE Trans. Automat. Contr.*, vol. 22, pp. 977-980, 1977.

[6] H. R. Sirisena, "Minimal-order observer for linear functions of the state vector," *Int. J. Contr.*, vol. 29, pp. 235-254, 1979.

[7] C-C. Tsui, "A new algorithm for the design of multifunctional observers," *IEEE Trans. Automat. Contr.*, vol. 30, pp. 89-93, 1985.

[8] C-C. Tsui, "On the order reduction of linear functional observers," *IEEE Trans. Automat. Contr.*, vol. 31, pp. 447-449, 1986.

[9] C-C. Tsui, "What is the minimum function observer order?, *Journal of the Franklin Institute* vol. 335, Issue 4, pp. 623-628, 1998.

[10] M. Aldeen and H. Trinh, "Reduced-order linear functional observer for linear systems," *IEE Proc.-Control Theory Appl.*, vol. 146, no. 5, pp. 399-405, 1999.

[11] C-C. Tsui, "What is the minimum function observer order?," *Proceeding of the 3rd World Congress on Intelligent Control and Automation*, pp. 2811-2816, Hefei, China, 2000.

[12] M. Darouach, "Existence and design of functional observers for linear systems," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 940-943, 2000.

[13] J. Moreno, Quasi-unknown input observers for linear systems, *Proceedings of the IEEE International Conference on Control Applications*, pp 732-737, 2001.

[14] C-C. Tsui, *Robust Control System Design, Advance State Space Techniques*, 2nd Edition, Taylor and Francis, 2004.

[15] J. Moreno, Simultaneous Observation of Linear Systems: A State-Space Interpretation, *IEEE Trans. Automat. Contr.*, vol. 50, pp. 1021-1025, 2005.

[16] H. Trinh, T. Fernando and S. Nahavandi, "Partial-state observers for nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 51, pp. 1808 - 1812, 2006.

[17] H. Trinh, T.D. Tran, and T. Fernando,"Disturbance Decoupled Observers for Systems With Unknown Inputs," *IEEE Trans. Automat. Contr.*, vol. 51, pp. 2397 - 2402, 2008.

[18] T.L. Fernando, H.M. Trinh and L.S. Jennings, "Functional observability and the design of minimum order functional observers," *IEEE Trans. Automat. Contr.*, vol. 55, pp. 1268 - 1273, 2010.