Aerospace Science and Technology 37 (2014) 70–80

Contents lists available at ScienceDirect

Aerospace Science and Technology

www.elsevier.com/locate/aescte

Unified nonlinear optimal flight control and state estimation of highly maneuverable aircraft

Ming Xin *, [1]

Department of Mechanical and Aerospace Engineering, University of Missouri, Columbia, MO 65211, United States

ARTICLE INFO

Article history:
Received 20 January 2014
Received in revised form 11 April 2014
Accepted 12 May 2014
Available online 16 May 2014

Keywords:
Flight control
Observer
State estimation
Optimal control

ABSTRACT

In this paper, a nonlinear flight control and a nonlinear state observer are designed in one unified framework for a high performance aircraft. The flight control is formulated as a nonlinear optimal control problem and the θ-D technique is utilized to design a closed-form optimal control law by solving the associated Hamilton–Jacobi–Bellman equation approximately via a perturbation process. Motivated by the notion that control and estimation are dual concepts, a new θ-D observer is developed to estimate the state used in the feedback by constructing the dual of the θ-D controller. Theoretical analysis of the proposed observer is given. Since both the optimal flight control and the state observer are unified in the same θ-D algorithm that leads to closed-form solutions to the optimal control and observer gains, this scheme greatly facilitates onboard implementation due to its computation efficiency. The unified design method is applied to a highly maneuverable aircraft operating at high angles of attack. Simulation results show that the θ-D control-θ-D observer suite exhibits excellent performance.

1. Introduction

Flight control of high performance aircraft at high angle of attack has received considerable interest due to its maneuverability advantage over the conventional aircraft [4,9]. However, the highly nonlinear dynamics presents great challenges in designing flight control systems. The traditional linear control methods are incapable of handling such challenges. A number of nonlinear control techniques have been extensively investigated in the recent decade to ensure adequate stability and performance at this extreme flight condition. Dynamic inversion [19] is one of the popular methods that offer great flexibility for handling changing models and nonstandard flight regimes. Wang and Stengel [24] employed a stochastic control technique to enhance the robustness of the dynamic inversion controller and improve the handling qualities of a high angle-of-attack aircraft. Neural adaptive control in [12,22] was combined with dynamic inversion to compensate for modeling and inversion errors. Nonlinear optimal control approaches have also been used for flight control of high performance aircraft. The state-dependent Riccati equation (SDRE) technique [3,15] has attracted a great deal of attention since it is a direct extension of the linear quadratic regulator (LQR) [2] formulation to nonlinear systems. It has a broad range of applications [3] including the aircraft flight control [1] and state estimation [16]. However, solving the SDRE on-line may impose an excessive computation load. In this paper, flight control of the highly maneuverable aircraft operating at high angle of attack is designed using the θ-D nonlinear optimal control technique [25]. The θ-D technique is a systematic nonlinear optimal control method derived from an approximate solution to the Hamilton–Jacobi–Bellman (HJB) equation using a perturbation approach. It provides a closed-form optimal control law and is easy to implement onboard. The main contribution of this paper is to propose a new θ-D observer for obtaining the states in the feedback by exploiting the dual concept between optimal control and observer. The optimal control and state estimation can be unified in the same θ-D algorithm to compute the closed-form feedback control law as well as the closed-form observer gain such that the design process can be significantly simplified.

Observer design for nonlinear systems has been an active field of research over the last few decades. Two primary approaches and their variations have been extensively studied. One well-established method is the exact error linearization [11] based on a nonlinear state transformation by which the error dynamics of state is linear so that the design of state observer can be performed using linear techniques. Many variations and generalizations to this approach include multi-input–multi-output error linearization [10,13], input derivative linearization [7], recursive backstepping-like nonlinear observer design [21], and dynamic observer error linearization by dynamic system extension and virtual

* Tel.: +1 573 882 7933.
E-mail address: xin@missouri.edu.
[1] Associate Professor.

http://dx.doi.org/10.1016/j.ast.2014.05.001
1270-9638/

outputs [17]. Although the error linearization approach has been extensively studied in theory, its integrability condition limits possible applications and the need to solve a system of partial differential equations (PDE) complicates the design process. The second approach to the nonlinear observer design involves high-gain techniques [18,23]. The system in consideration is assumed to possess a structure with a linear and nonlinear part. A linear observer is designed such that the linear part dominates the nonlinearity. The Lipschitz condition is assumed to hold for the nonlinear part and the Lipschitz constant is needed to determine the eigenvalues of the error dynamics for stability. Thau's method [23] is literally a verification method, rather than a design technique since finding an observer gain to satisfy the Thau's condition is not easy. Rajamani [18] gave necessary and sufficient existence conditions for high gain observers and shows that merely placing the eigenvalues far into the left half-plane is not sufficient to guarantee stability. Associated matrix of eigenvectors must be well-conditioned. The iterative algorithm presented in this paper for computing the observer gains is quite involved. Zhu and Han [28] extended these conclusions by designing a reduced-order observer based on the solution of a Riccati equation. Robenack and Lynch [20] proposed two observers combining the exact error linearization and the high-gain observer result. The resultant observer gain can be computed efficiently without knowledge of either the observer canonical form or the associated transformation. However, the requirements of existence of the observer canonical forms, the integrability condition, and the single output measurement, limit the application of this observer to only a restricted class of systems.

Extended Kalman Filter (EKF) [6] and its variants are widely used to estimate the states in stochastic settings. Since linearization about the current estimate is used in each step to compute the observer gain, these methods cannot offer stability guarantees and/or robustness against modeling errors. The State Dependent Riccati Equation Filter (SDREF) [8,16] has received much attention during the last decade. It is a dual version of the SDRE controller [3] and does not need any linearization like the EKF. Stability of the SDREF has not been theoretically established and on-line solution to the algebraic Riccati equation at each instant, is computationally intensive.

In this paper, a new nonlinear observer, called the θ-D observer is developed based on the dual of the θ-D controller [25]. The θ-D observer takes the same structure as the continuous steady-state linear Kalman filter *but* it is applicable to nonlinear estimation problems. It is different from the EKF in that no linearization is required. It is also similar in structure to the SDREF. However, the θ-D observer provides *closed-form* observer gains and thus obviates the need for time-wise costly on-line numerical solutions of the algebraic Riccati equation. Compared with the exact error linearization and high gain observers [20], the θ-D observer can be applied to a much broader class of nonlinear systems because it does not requires the system to be transformable to the canonical form and the integrability condition. The theoretical proofs on convergence and stability of this new observer are given. Combining the θ-D optimal control and the θ-D observer in one unified framework greatly facilitates the design process and implementation since they share the same θ-D algorithm. The excellent performance of the unified scheme is demonstrated by the flight control design of a highly maneuverable aircraft.

In Section 2, the θ-D optimal control technique is reviewed. Development of the θ-D observer and associated theoretical results are given in Section 3. Applications to the flight control problem and simulation results are presented in Section 4. Some concluding remarks are given in Section 5.

2. Review of θ-D control technique

The θ-D nonlinear control technique addresses the class of nonlinear time-invariant systems described by

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{u} \tag{1}$$

with the cost functional:

$$J = \frac{1}{2}\int_0^\infty [\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{u}^T\boldsymbol{R}\boldsymbol{u}]dt \tag{2}$$

where $\boldsymbol{x} \in \Omega \subset \mathrm{R}^n$, $\boldsymbol{f} : \Omega \to \mathrm{R}^n$, $\boldsymbol{B} \in \mathrm{R}^{n\times m}$, $\boldsymbol{u} : \Omega \to \mathrm{R}^m$, $\boldsymbol{Q} \in \mathrm{R}^{n\times n}$, $\boldsymbol{R} \in \mathrm{R}^{m\times m}$; Ω is a compact subset in R^n; $\boldsymbol{Q}$ is a positive semi-definite matrix and $\boldsymbol{R}$ is a positive definite matrix; $\boldsymbol{B}$ is a constant matrix and $\boldsymbol{f}(\boldsymbol{0}) = \boldsymbol{0}$. It is also assumed that $\boldsymbol{f}(\boldsymbol{x})$ is continuously differentiable and zero state observable through $\boldsymbol{Q}$.

The θ-D control method can be summarized as the following procedure [25].

Write the original nonlinear state equation as a linear-like structure:

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{u} = \boldsymbol{F}(\boldsymbol{x})\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u} = \left[\boldsymbol{A}_0 + \theta\frac{\boldsymbol{A}(\boldsymbol{x})}{\theta}\right]\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u} \tag{3}$$

where $\boldsymbol{A}_0$ is a constant matrix such that $(\boldsymbol{A}_0, \boldsymbol{B})$ is a controllable pair and $[\boldsymbol{F}(\boldsymbol{x}), \boldsymbol{B}]$ is pointwise controllable. θ is an intermediate variable for the convenience of power series expansion that will be used in the optimal control approximation.

By virtue of a perturbation process [25], a suboptimal feedback controller can be obtained by

$$\boldsymbol{u} = -\boldsymbol{R}^{-1}\boldsymbol{B}^T\sum_{i=0}^{\infty}\boldsymbol{T}_i(\boldsymbol{x},\theta)\theta^i\boldsymbol{x} \quad (i = 0, \cdots, n, \cdots) \tag{4}$$

where $\boldsymbol{T}_i(\boldsymbol{x},\theta)$ is a symmetric matrix and is obtained by recursively solving the following equations:

$$\boldsymbol{T}_0\boldsymbol{A}_0 + \boldsymbol{A}_0^T\boldsymbol{T}_0 - \boldsymbol{T}_0\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T\boldsymbol{T}_0 + \boldsymbol{Q} = \boldsymbol{0} \tag{5a}$$

$$\begin{aligned}&\boldsymbol{T}_1\left(\boldsymbol{A}_0 - \boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T\boldsymbol{T}_0\right) + \left(\boldsymbol{A}_0^T - \boldsymbol{T}_0\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T\right)\boldsymbol{T}_1\\&\quad = -\frac{\boldsymbol{T}_0\boldsymbol{A}(\boldsymbol{x})}{\theta} - \frac{\boldsymbol{A}^T(\boldsymbol{x})\boldsymbol{T}_0}{\theta} - \boldsymbol{D}_1\end{aligned} \tag{5b}$$

$$\begin{aligned}&\boldsymbol{T}_2\left(\boldsymbol{A}_0 - \boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T\boldsymbol{T}_0\right) + \left(\boldsymbol{A}_0^T - \boldsymbol{T}_0\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T\right)\boldsymbol{T}_2\\&\quad = -\frac{\boldsymbol{T}_1\boldsymbol{A}(\boldsymbol{x})}{\theta} - \frac{\boldsymbol{A}^T(\boldsymbol{x})\boldsymbol{T}_1}{\theta} + \boldsymbol{T}_1\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T\boldsymbol{T}_1 - \boldsymbol{D}_2\end{aligned} \tag{5c}$$

$$\vdots$$

$$\begin{aligned}&\boldsymbol{T}_n\left(\boldsymbol{A}_0 - \boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T\boldsymbol{T}_0\right) + \left(\boldsymbol{A}_0^T - \boldsymbol{T}_0\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T\right)\boldsymbol{T}_n\\&\quad = -\frac{\boldsymbol{T}_{n-1}\boldsymbol{A}(\boldsymbol{x})}{\theta} - \frac{\boldsymbol{A}^T(\boldsymbol{x})\boldsymbol{T}_{n-1}}{\theta}\\&\qquad + \sum_{j=1}^{n-1}\boldsymbol{T}_j\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T\boldsymbol{T}_{n-j} - \boldsymbol{D}_n\end{aligned} \tag{5d}$$

It is easy to observe that Eq. (5a) is an algebraic Riccati equation. Under the controllability and observability conditions, the solution $\boldsymbol{T}_0$ will be a positive definite constant matrix. Eqs. (5b)–(5d) are Lyapunov equations that are *linear* in terms of $\boldsymbol{T}_i$ $(i = 1, \cdots, n)$. Since all the coefficients of $\boldsymbol{T}_i$ $(i = 1, \cdots, n)$ on the left-hand side of the equations are the same constant matrices, i.e. $\boldsymbol{A}_0 - \boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T\boldsymbol{T}_0$ and $\boldsymbol{A}_0^T - \boldsymbol{T}_0\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T$, closed-form solution for $\boldsymbol{T}_i(\boldsymbol{x},\theta)$ can be easily obtained by solving Eqs. (5b)–(5d) successively with some linear algebra [25], which will also be discussed in detail in the next section, θ-D observer design.

The $\boldsymbol{D}_i$ matrix is constructed in the form of:

$$\boldsymbol{D}_1 = k_1 e^{-l_1 t}\left[-\frac{\boldsymbol{T}_0\boldsymbol{A}(\boldsymbol{x})}{\theta} - \frac{\boldsymbol{A}^T(\boldsymbol{x})\boldsymbol{T}_0}{\theta}\right] \tag{6a}$$

$$\boldsymbol{D}_2 = k_2 e^{-l_2 t}\left[-\frac{\boldsymbol{T}_1\boldsymbol{A}(\boldsymbol{x})}{\theta} - \frac{\boldsymbol{A}^T(\boldsymbol{x})\boldsymbol{T}_1}{\theta} + \boldsymbol{T}_1\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T\boldsymbol{T}_1\right] \tag{6b}$$

$$\vdots$$

$$\boldsymbol{D}_n = k_n e^{-l_n t}\left[-\frac{\boldsymbol{T}_{n-1}\boldsymbol{A}(\boldsymbol{x})}{\theta} - \frac{\boldsymbol{A}^T(\boldsymbol{x})\boldsymbol{T}_{n-1}}{\theta} + \sum_{j=1}^{n-1}\boldsymbol{T}_j\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T\boldsymbol{T}_{n-j}\right] \tag{6c}$$

where k_i and $l_i > 0$ are design parameters to modulate transient performance. $\boldsymbol{D}_i$ is chosen such that

$$-\frac{\boldsymbol{T}_{i-1}\boldsymbol{A}(\boldsymbol{x})}{\theta} - \frac{\boldsymbol{A}^T(\boldsymbol{x})\boldsymbol{T}_{i-1}}{\theta} + \sum_{j=1}^{i-1}\boldsymbol{T}_j\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T\boldsymbol{T}_{i-j} - \boldsymbol{D}_i = \varepsilon_i\left[-\frac{\boldsymbol{T}_{i-1}\boldsymbol{A}(\boldsymbol{x})}{\theta} - \frac{\boldsymbol{A}^T(\boldsymbol{x})\boldsymbol{T}_{i-1}}{\theta} + \sum_{j=1}^{i-1}\boldsymbol{T}_j\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T\boldsymbol{T}_{i-j}\right] \tag{7}$$

where $\varepsilon_i = 1 - k_i e^{-l_i t}$ $(i = 1, \cdots n)$ is a small number, i.e. $0 \le \varepsilon_i \le 1$. ε_i is chosen to be a small number in order to overcome the large control problem due to the state dependent term $\boldsymbol{A}(\boldsymbol{x})$ in Eqs. (5b)–(5d). ε_i is also required in the proof of convergence and stability of the above algorithm [25].

Remark 2.1. θ is merely an intermediate variable. The introduction of θ is for the convenience of power series expansion, and it is canceled when $\boldsymbol{T}_i(\boldsymbol{x}, \theta)$ multiplies θ^i in the final control calculations, i.e., Eq. (4). The cancellation will be more clearly seen in the convergence proof of the θ-D observer design in the next section.

The θ-D optimal control technique has been successfully applied to many practical engineering problems, e.g. Refs. [26,27].

3. Formulation of θ-D observer

Motivated by the duality property between the linear optimal regulator and observer, we can formulate the θ-D observer as the counterpart of the θ-D controller.

Consider the nonlinear system and measurements given respectively by

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) \tag{8}$$

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} \tag{9}$$

Note that the feedback control is included in $\boldsymbol{f}(\boldsymbol{x})$. $\boldsymbol{f}(\boldsymbol{x})$ is assumed to be of class C^1, and $\boldsymbol{H}$ is a constant matrix that relates the states $\boldsymbol{x}$ to the observations denoted by $\boldsymbol{y}$. Rewrite (8) as:

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{F}(\boldsymbol{x})\boldsymbol{x} = \left[\boldsymbol{A}_0 + \theta\frac{\boldsymbol{A}(\boldsymbol{x})}{\theta}\right]\boldsymbol{x} \tag{10}$$

The θ-D observer is constructed as

$$\dot{\hat{\boldsymbol{x}}} = \boldsymbol{F}(\hat{\boldsymbol{x}})\hat{\boldsymbol{x}} + \boldsymbol{K}_f(\hat{\boldsymbol{x}})\left[\boldsymbol{y}(\boldsymbol{x}) - \boldsymbol{H}\hat{\boldsymbol{x}}\right] = \left[\boldsymbol{A}_0 + \theta\frac{\boldsymbol{A}(\hat{\boldsymbol{x}})}{\theta}\right]\hat{\boldsymbol{x}} + \boldsymbol{K}_f(\hat{\boldsymbol{x}})\left[\boldsymbol{y}(\boldsymbol{x}) - \boldsymbol{H}\hat{\boldsymbol{x}}\right] \tag{11}$$

where

$$\boldsymbol{K}_f(\hat{\boldsymbol{x}}) = \boldsymbol{P}(\hat{\boldsymbol{x}})\boldsymbol{H}^T\boldsymbol{V}^{-1} \tag{12}$$

$$\boldsymbol{P}(\hat{\boldsymbol{x}}) = \sum_{i=0}^{\infty}\hat{\boldsymbol{T}}_i(\hat{\boldsymbol{x}}, \theta)\theta^i \tag{13}$$

and $\hat{\boldsymbol{T}}_i(\hat{\boldsymbol{x}}, \theta)$ is the solution to the following equations in which we take the dual of $\boldsymbol{A}_0$, $\boldsymbol{A}(\boldsymbol{x})$ and $\boldsymbol{B}$ in (5a)–(5d), i.e. $\boldsymbol{A}_0^T$, $\boldsymbol{A}^T(\hat{\boldsymbol{x}})$ and $\boldsymbol{H}^T$.

$$\hat{\boldsymbol{T}}_0\boldsymbol{A}_0^T + \boldsymbol{A}_0\hat{\boldsymbol{T}}_0 - \hat{\boldsymbol{T}}_0\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_0 + \boldsymbol{W} = \boldsymbol{0} \tag{14a}$$

$$\hat{\boldsymbol{T}}_1\left(\boldsymbol{A}_0^T - \boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_0\right) + \left(\boldsymbol{A}_0 - \hat{\boldsymbol{T}}_0\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\right)\hat{\boldsymbol{T}}_1 = -\frac{\hat{\boldsymbol{T}}_0\boldsymbol{A}^T(\hat{\boldsymbol{x}})}{\theta} - \frac{\boldsymbol{A}(\hat{\boldsymbol{x}})\hat{\boldsymbol{T}}_0}{\theta} - \hat{\boldsymbol{D}}_1 \tag{14b}$$

$$\hat{\boldsymbol{T}}_2\left(\boldsymbol{A}_0^T - \boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_0\right) + \left(\boldsymbol{A}_0 - \hat{\boldsymbol{T}}_0\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\right)\hat{\boldsymbol{T}}_2 = -\frac{\hat{\boldsymbol{T}}_1\boldsymbol{A}^T(\hat{\boldsymbol{x}})}{\theta} - \frac{\boldsymbol{A}(\hat{\boldsymbol{x}})\hat{\boldsymbol{T}}_1}{\theta} + \hat{\boldsymbol{T}}_1\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_1 - \hat{\boldsymbol{D}}_2 \tag{14c}$$

$$\vdots$$

$$\hat{\boldsymbol{T}}_n\left(\boldsymbol{A}_0^T - \boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_0\right) + \left(\boldsymbol{A}_0 - \hat{\boldsymbol{T}}_0\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\right)\hat{\boldsymbol{T}}_n = -\frac{\hat{\boldsymbol{T}}_{n-1}\boldsymbol{A}^T(\hat{\boldsymbol{x}})}{\theta} - \frac{\boldsymbol{A}(\hat{\boldsymbol{x}})\hat{\boldsymbol{T}}_{n-1}}{\theta} + \sum_{j=1}^{n-1}\hat{\boldsymbol{T}}_j\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_{n-j} - \hat{\boldsymbol{D}}_n \tag{14d}$$

where $\hat{\boldsymbol{D}}_1, \cdots, \hat{\boldsymbol{D}}_n$ have the similar expressions as (6),

$$\hat{\boldsymbol{D}}_1 = k_1 e^{-l_1 t}\left[-\frac{\hat{\boldsymbol{T}}_0\boldsymbol{A}^T(\hat{\boldsymbol{x}})}{\theta} - \frac{\boldsymbol{A}(\hat{\boldsymbol{x}})\hat{\boldsymbol{T}}_0}{\theta}\right] \tag{15a}$$

$$\hat{\boldsymbol{D}}_2 = k_2 e^{-l_2 t}\left[-\frac{\hat{\boldsymbol{T}}_1\boldsymbol{A}^T(\hat{x})}{\theta} - \frac{\boldsymbol{A}(\hat{x})\hat{\boldsymbol{T}}_1}{\theta} + \hat{\boldsymbol{T}}_1\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_1\right] \tag{15b}$$

$$\vdots$$

$$\hat{\boldsymbol{D}}_n = k_n e^{-l_n t}\left[-\frac{\hat{\boldsymbol{T}}_{n-1}\boldsymbol{A}^T(\hat{\boldsymbol{x}})}{\theta} - \frac{\boldsymbol{A}(\hat{\boldsymbol{x}})\hat{\boldsymbol{T}}_{n-1}}{\theta} + \sum_{j=1}^{n-1}\hat{\boldsymbol{T}}_j\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_{n-j}\right] \tag{15c}$$

such that

$$-\frac{\hat{\boldsymbol{T}}_{i-1}\boldsymbol{A}(\hat{\boldsymbol{x}})}{\theta} - \frac{\boldsymbol{A}^T(\hat{\boldsymbol{x}})\hat{\boldsymbol{T}}_{i-1}}{\theta} + \sum_{j=1}^{i-1}\hat{\boldsymbol{T}}_j\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_{i-j} - \hat{\boldsymbol{D}}_i = \varepsilon_i\left[-\frac{\hat{\boldsymbol{T}}_{i-1}\boldsymbol{A}^T(\hat{\boldsymbol{x}})}{\theta} - \frac{\boldsymbol{A}(\hat{\boldsymbol{x}})\hat{\boldsymbol{T}}_{i-1}}{\theta} + \sum_{j=1}^{i-1}\hat{\boldsymbol{T}}_j\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_{i-j}\right]$$

where $\varepsilon_i = 1 - k_i e^{-l_i t}$ $(i = 1, \cdots n)$ is a small number, i.e. $0 \le \varepsilon_i \le 1$.

In the above equations, $\boldsymbol{V} > 0$ and $\boldsymbol{W} \ge 0$ are treated as weights to improve the convergence of the observer. Note that they can be construed as noise covariance matrices in a stochastic setting although they do not have statistics based interpretations in a deterministic observer design.

In order to obtain the θ-D observer gain, Eqs. (14a)–(14d) are solved recursively. The steps are summarized as follows:

1) Solve the algebraic Riccati equation (14a) to obtain $\hat{\boldsymbol{T}}_0$ once $\boldsymbol{A}_0$, $\boldsymbol{H}$, $\boldsymbol{W}$ and $\boldsymbol{V}$ are determined. Note that the resulting $\hat{\boldsymbol{T}}_0$ is a positive-definite constant matrix.

2) Solve the Lyapunov equation (14b) to obtain $\hat{\boldsymbol{T}}_1(\hat{\boldsymbol{x}}, \theta)$. Note that this is a linear equation in terms of $\hat{\boldsymbol{T}}_1(\hat{\boldsymbol{x}}, \theta)$ and a unique

property of this equation and Eqs. (14c) and (14d) is that the coefficient matrices $\boldsymbol{A}_0^T - \boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_0$ and $\boldsymbol{A}_0 - \hat{\boldsymbol{T}}_0\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}$ are constant matrices. Assume that $\boldsymbol{A}_{c_0} = \boldsymbol{A}_0 - \hat{\boldsymbol{T}}_0\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}$. Through linear algebra, Eq. (14b) can be brought into the form $\hat{\boldsymbol{A}}_0 \operatorname{vec}(\hat{\boldsymbol{T}}_1) = \operatorname{vec}[\boldsymbol{Q}_1(\hat{\boldsymbol{x}},\theta,t)]$ where $\boldsymbol{Q}_1(\hat{\boldsymbol{x}},\theta,t)$ contains all the nonlinear state dependent terms on the right-hand side of Eq. (14b); $\operatorname{vec}(\hat{\boldsymbol{T}}_1)$ and $\operatorname{vec}[\boldsymbol{Q}_1(\hat{\boldsymbol{x}},\theta,t)]$ denote stacking the elements of the matrix by rows in a vector form; $\hat{\boldsymbol{A}}_0 = \boldsymbol{I}_n \otimes \boldsymbol{A}_{c_0}^{\mathrm{T}} + \boldsymbol{A}_{c_0} \otimes \boldsymbol{I}_n$ is a constant matrix and the symbol $\otimes$ denotes Kronecker product. Thus, the resulting solution of $\hat{\boldsymbol{T}}_1$ can be written in a closed-form expression $\operatorname{vec}(\hat{\boldsymbol{T}}_1) = \hat{\boldsymbol{A}}_0^{-1} \operatorname{vec}[\boldsymbol{Q}_1(\hat{\boldsymbol{x}},\theta,t)]$.

3) Solve Eqs. (14c) and (14d) for $\hat{\boldsymbol{T}}_2(\hat{\boldsymbol{x}},\theta)$ and $\hat{\boldsymbol{T}}_n(\hat{\boldsymbol{x}},\theta)$ following the similar procedure in Step 2.

Since all the coefficients of $\hat{\boldsymbol{T}}_i, i = 1,\cdots n$ on the left-hand side of Eqs. (14a)–(14d) are the same constant matrices, i.e. $\boldsymbol{A}_0^T - \boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_0$ and $\boldsymbol{A}_0 - \hat{\boldsymbol{T}}_0\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}$, closed-form solution for all $\hat{\boldsymbol{T}}_i(\hat{\boldsymbol{x}},\theta)$ can be easily obtained with just one matrix inverse operation, i.e. $\hat{\boldsymbol{A}}_0^{-1}$. Thus, we can get the closed-form observer gain $\boldsymbol{P}(\hat{\boldsymbol{x}}) = \sum_{i=0}^{\infty} \hat{\boldsymbol{T}}_i(\hat{\boldsymbol{x}},\theta)\theta^i$ if we take a finite number of terms. Usually the first three terms $\hat{\boldsymbol{T}}_0, \hat{\boldsymbol{T}}_1$ and $\hat{\boldsymbol{T}}_2$ in the observer equation (13) are sufficient to achieve satisfactory performance.

The following theorem shows the convergence of the series $\sum_{i=0}^{\infty} \hat{\boldsymbol{T}}_i(\hat{\boldsymbol{x}},\theta)\theta^i$.

Theorem 3.1. *If the following conditions are satisfied*:

(i) $\hat{\boldsymbol{x}} \in \Omega$, *where* $\Omega \subset \mathrm{R}^n$ *is a compact set*;
(ii) $(\boldsymbol{A}_0, \boldsymbol{H})$ *and* $(\boldsymbol{A}_0^T, \boldsymbol{G})$ *are observable, where* $\boldsymbol{W} = \boldsymbol{G}\boldsymbol{G}^T$;
(iii) $\boldsymbol{A}(\hat{\boldsymbol{x}})$ *is continuous on* Ω *and* $\|\boldsymbol{A}(\hat{\boldsymbol{x}})\|_2 \neq 0, \forall \hat{\boldsymbol{x}} \in \Omega$;
(iv) $\lambda_{\max}[(\boldsymbol{A}_0 - \hat{\boldsymbol{T}}_0\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}) + (\boldsymbol{A}_0^T - \boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_0)] < 0$, *where* $\lambda_{\max}$ *denotes the largest eigenvalue, the series* $\sum_{i=0}^{\infty} \hat{\boldsymbol{T}}_i(\hat{\boldsymbol{x}},\theta)\theta^i$ *obtained by the algorithm* (14a)–(14d) *is pointwise convergent.*

Proof. Considering (14b) and the selection of $\hat{\boldsymbol{D}}_1$ in (15a), Eq. (14b) can be written as:

$$\hat{\boldsymbol{T}}_1(\boldsymbol{A}_0^T - \boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_0) + (\boldsymbol{A}_0 - \hat{\boldsymbol{T}}_0\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H})\hat{\boldsymbol{T}}_1 = -\varepsilon_1(\hat{\boldsymbol{T}}_0\boldsymbol{A}^T + \boldsymbol{A}\hat{\boldsymbol{T}}_0)\frac{1}{\theta} \tag{16}$$

with

$$\varepsilon_1 = 1 - k_1 e^{-l_1 t} \tag{17}$$

For brevity, we omit the argument $\hat{\boldsymbol{x}}$ in $\boldsymbol{A}(\hat{\boldsymbol{x}})$ and t in $\varepsilon_i(t)$ to simplify the notation. Assume that the solution to the equation

$$\bar{\boldsymbol{T}}_1(\boldsymbol{A}_0^T - \boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_0) + (\boldsymbol{A}_0 - \hat{\boldsymbol{T}}_0\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H})\bar{\boldsymbol{T}}_1 = -\varepsilon_1(\hat{\boldsymbol{T}}_0\boldsymbol{A}^T + \boldsymbol{A}\hat{\boldsymbol{T}}_0) \tag{18}$$

is $\bar{\boldsymbol{T}}_1$. Also assume that

$$\bar{\boldsymbol{T}}_0 = \hat{\boldsymbol{T}}_0 \tag{19}$$

Using the linearity of the Lyapunov equation (18), the solution to (16) becomes

$$\hat{\boldsymbol{T}}_1 = \frac{1}{\theta}\bar{\boldsymbol{T}}_1 \tag{20}$$

Similarly assume that the solution to the equation

$$\bar{\boldsymbol{T}}_2(\boldsymbol{A}_0^T - \boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_0) + (\boldsymbol{A}_0 - \hat{\boldsymbol{T}}_0\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H})\bar{\boldsymbol{T}}_2 = -\varepsilon_2(\bar{\boldsymbol{T}}_1\boldsymbol{A}^T + \boldsymbol{A}\bar{\boldsymbol{T}}_1 - \bar{\boldsymbol{T}}_1\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\bar{\boldsymbol{T}}_1) \tag{21}$$

is $\bar{\boldsymbol{T}}_2$. Then the solution to (14c) using (15b) and (20) becomes:

$$\hat{\boldsymbol{T}}_2 = \frac{1}{\theta^2}\bar{\boldsymbol{T}}_2 \tag{22}$$

In the same manner,

$$\hat{\boldsymbol{T}}_n = \frac{1}{\theta^n}\bar{\boldsymbol{T}}_n \tag{23}$$

where $\bar{\boldsymbol{T}}_n$ is the solution to

$$\bar{\boldsymbol{T}}_n(\boldsymbol{A}_0^T - \boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_0) + (\boldsymbol{A}_0 - \hat{\boldsymbol{T}}_0\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H})\bar{\boldsymbol{T}}_n = -\varepsilon_n\left[\bar{\boldsymbol{T}}_{n-1}\boldsymbol{A}^T + \boldsymbol{A}\bar{\boldsymbol{T}}_{n-1} - \sum_{j=1}^{n-1}\bar{\boldsymbol{T}}_j\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\bar{\boldsymbol{T}}_{n-j}\right] \tag{24}$$

and

$$\varepsilon_n = 1 - k_n e^{-l_n t}. \tag{25}$$

From (19), (20), (22), and (23) we note that proving the convergence of $\sum_{i=0}^{\infty} \hat{\boldsymbol{T}}_i(\hat{\boldsymbol{x}},\theta)\theta^i$ is equivalent to proving the convergence of $\sum_{i=0}^{\infty} \bar{\boldsymbol{T}}_i(\hat{\boldsymbol{x}})$ where $\bar{\boldsymbol{T}}_i$ satisfy Eqs. (14a), (18), (21) and (24) because θ^i gets canceled.

Note that this part substantiates Remark 2.1 that θ^i is canceled in the control and observer gain calculations.

In order to prove the convergence of the series $\sum_{i=0}^{\infty} \bar{\boldsymbol{T}}_i(\hat{\boldsymbol{x}})$, one needs to find a norm bound for each $\bar{\boldsymbol{T}}_i$.

Given a Lyapunov equation

$$\bar{\boldsymbol{A}}^T\bar{\boldsymbol{P}} + \bar{\boldsymbol{P}}\bar{A} = -\bar{\boldsymbol{Q}} \tag{26}$$

where $\bar{\boldsymbol{A}}, \bar{\boldsymbol{P}}, \bar{\boldsymbol{Q}} \in \mathrm{R}^{n\times n}$, we have the norm bound for $\bar{\boldsymbol{P}}$ [14]

$$\|\bar{\boldsymbol{P}}\|_\bullet \le \frac{\|\bar{\boldsymbol{Q}}\|_\bullet}{-\mu_\bullet(\bar{\boldsymbol{A}}^T) - \mu_\bullet(\bar{\boldsymbol{A}})} \tag{27}$$

if $\bar{\boldsymbol{A}}$ is a Hurwitz matrix and $-\mu_\bullet(\bar{\boldsymbol{A}}^T) - \mu_\bullet(\bar{\boldsymbol{A}}) > 0$ where $\mu_\bullet(\bar{\boldsymbol{A}})$ is a matrix measure of $\bar{\boldsymbol{A}}$ induced from $\|\bullet\|_\bullet$. In the case of 2-norm,

$$\mu_2(\bar{\boldsymbol{A}}) \triangleq \frac{1}{2}\lambda_{\max}(\bar{\boldsymbol{A}} + \bar{\boldsymbol{A}}^T) \tag{28}$$

In the following, $\|\cdot\|$ is defined as a 2-norm and denotes $\mu(\cdot) = \mu_2(\cdot)$.

Since $(\boldsymbol{A}_0, \boldsymbol{H})$ is an observable pair, $(\boldsymbol{A}_0^T, \boldsymbol{H}^T)$ is controllable. So, condition (ii) implies that the Riccati equation (14a) has a positive definite solution $\hat{\boldsymbol{T}}_0$ (note $\hat{\boldsymbol{T}}_0 = \bar{\boldsymbol{T}}_0$) and $\boldsymbol{A}_0^T - \boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_0$ is a Hurwitz matrix. Given this result and the condition (iv), the norm bound for $\bar{\boldsymbol{T}}_1$ can be derived from Eq. (18) and inequality (27):

$$\|\bar{\boldsymbol{T}}_1\| \le \frac{\|\varepsilon_1[\hat{\boldsymbol{T}}_0\boldsymbol{A}^T + \boldsymbol{A}\hat{\boldsymbol{T}}_0]\|}{-\mu(\boldsymbol{A}_0 - \hat{\boldsymbol{T}}_0\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}) - \mu(\boldsymbol{A}_0^T - \boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_0)} \tag{29}$$

Let

$$C = \frac{1}{-\mu(\boldsymbol{A}_0 - \hat{\boldsymbol{T}}_0\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}) - \mu(\boldsymbol{A}_0^T - \boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{T}}_0)} \tag{30}$$

Using the condition (iv), we have $C > 0$.

Then

$$\|\bar{\boldsymbol{T}}_1\| \le C\varepsilon_1\|\hat{\boldsymbol{T}}_0\boldsymbol{A}^T + \boldsymbol{A}\hat{\boldsymbol{T}}_0\| \le C\varepsilon_1[\|\hat{\boldsymbol{T}}_0\|(\|\boldsymbol{A}\| + \|\boldsymbol{A}^T\|)] \tag{31}$$

Since $\boldsymbol{A}(\hat{\boldsymbol{x}})$ is continuous (condition (iii)) on a compact set Ω, it is bounded on Ω.

Let

$$C_1 = \max_{\hat{\boldsymbol{x}}\in\Omega}(\|\boldsymbol{A}(\hat{\boldsymbol{x}})\| + \|\boldsymbol{A}^T(\hat{\boldsymbol{x}})\|) \tag{32}$$

Then

$$\|\bar{\boldsymbol{T}}_1\| \le \varepsilon_1 C C_1 \|\hat{\boldsymbol{T}}_0\|. \tag{33}$$

Condition (iii) implies that $C_1 \neq 0$. If it is zero, the nonlinear system will reduce to the linear system for which the solution is $\hat{\boldsymbol{T}}_0$, the solution to the Riccati equation (14a).

For later use, define

$$S_0 = \|\bar{\boldsymbol{T}}_0\| = \|\hat{\boldsymbol{T}}_0\|, \tag{34}$$

and

$$S_1 = \varepsilon_1 C C_1 \|\hat{\boldsymbol{T}}_0\| \tag{35}$$

Then

$$S_1 = O(\varepsilon_1). \tag{36}$$

Therefore, by choosing a sufficiently small ε_1, we can always make $\frac{S_1}{S_0} = \varepsilon_1 C C_1 < 1$.

Consider Eq. (21). A norm-bounded inequality for $\bar{\boldsymbol{T}}_2$ becomes

$$\|\bar{\boldsymbol{T}}_2\| \le C\varepsilon_2 \left\| \bar{\boldsymbol{T}}_1 \boldsymbol{A}^T + \boldsymbol{A}\bar{\boldsymbol{T}}_1 - \bar{\boldsymbol{T}}_1 \boldsymbol{H}^T \boldsymbol{V}^{-1} \boldsymbol{H} \bar{\boldsymbol{T}}_1 \right\| \tag{37}$$

Let $C_H \triangleq \|\boldsymbol{H}^T \boldsymbol{V}^{-1} \boldsymbol{H}\|$, a constant. Then

$$\begin{aligned} \|\bar{\boldsymbol{T}}_2\| &\le C \cdot \varepsilon_2 \left(C_1 \|\bar{\boldsymbol{T}}_1\| + \|\bar{\boldsymbol{T}}_1\|^2 C_H \right) \\ &\le C \cdot \varepsilon_2 \left(C_1 \cdot \varepsilon_1 C C_1 \|\hat{\boldsymbol{T}}_0\| + \varepsilon_1^2 C^2 C_1^2 \|\hat{\boldsymbol{T}}_0\|^2 C_H \right) \\ &= C^2 \varepsilon_1 \varepsilon_2 C_1^2 \|\hat{\boldsymbol{T}}_0\| \left(1 + \varepsilon_1 C \|\hat{\boldsymbol{T}}_0\| C_H \right) \end{aligned} \tag{38}$$

Let

$$C_2 = \max_{t \in [0,T]} \left(1 + \varepsilon_1 C \|\hat{\boldsymbol{T}}_0\| C_H \right). \tag{39}$$

It leads to

$$\|\bar{\boldsymbol{T}}_2\| \le \varepsilon_1 \varepsilon_2 C^2 \cdot C_1^2 \cdot C_2 \|\hat{\boldsymbol{T}}_0\| \tag{40}$$

Let

$$S_2 = \varepsilon_1 \varepsilon_2 C^2 \cdot C_1^2 \cdot C_2 \|\hat{\boldsymbol{T}}_0\| \tag{41}$$

which implies that

$$S_2 = O(\varepsilon_1 \varepsilon_2). \tag{42}$$

Note that from Eq. (35) it can be deduced that

$$\frac{S_2}{S_1} = \varepsilon_2 C C_1 C_2 = O(\varepsilon_2) \tag{43}$$

Therefore, if ε_2 is chosen sufficiently small, we can make

$$\frac{S_2}{S_1} < 1 \tag{44}$$

In a similar manner we can derive for $\bar{\boldsymbol{T}}_n$ $(n > 2)$ that

$$\|\bar{\boldsymbol{T}}_n\| \le (\varepsilon_1 \cdots \varepsilon_n) \cdot C^n C_1^n C_2 \ldots C_n \|\hat{\boldsymbol{T}}_0\| \tag{45}$$

and

$$(\varepsilon_1 \cdots \varepsilon_n) \cdot C^n C_1^n C_2 \ldots C_n \|\hat{\boldsymbol{T}}_0\| = O(\varepsilon_1 \cdots \varepsilon_n) \tag{46}$$

Once the bound for each $\bar{\boldsymbol{T}}_i$ is determined, the convergence of the series $\sum_{i=0}^{\infty} \bar{\boldsymbol{T}}_i$ can be proved as follows.

Define a series $\sum_{n=0}^{\infty} S_n$ with S_0 and S_1 defined in (34) and (35) and

$$S_n = (\varepsilon_1 \cdots \varepsilon_n) \cdot C^n C_1^n C_2 \ldots C_n \|\hat{\boldsymbol{T}}_0\| \tag{47}$$

Then

$$\frac{S_n}{S_{n-1}} = \varepsilon_n \cdot C C_1 C_n = O(\varepsilon_n) \tag{48}$$

By choosing a sufficiently small ε_n such that $\lim_{n\to\infty} \varepsilon_n \cdot C C_1 C_n < 1$, $\sum_{i=0}^{\infty} S_i$ becomes a convergent series. Since each $\|\bar{\boldsymbol{T}}_i\| \le S_i$, $\sum_{i=0}^{\infty} \bar{\boldsymbol{T}}_i$ is also a convergent series. Thus $\sum_{i=0}^{\infty} \hat{\boldsymbol{T}}_i(\hat{\boldsymbol{x}}, \theta)\theta^i$ is convergent. □

The following lemma proves the asymptotic stability of the θ-D observer.

Lemma 3.1. *Suppose that $\boldsymbol{f}_1(\boldsymbol{x}) = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{x}$ is Lipschitz continuous on a compact set $\Omega \subset \mathrm{R}^n$ with $\boldsymbol{x}, \hat{\boldsymbol{x}} \in \Omega$ and the conditions in Theorem 3.1 are satisfied, then the error dynamics defined by $\boldsymbol{e} = \hat{\boldsymbol{x}} - \boldsymbol{x}$ is asymptotically stable.*

Proof. Rewrite Eqs. (10) and (11) as

$$\dot{\boldsymbol{x}} = \boldsymbol{A}_0 \boldsymbol{x} + \boldsymbol{f}_1(\boldsymbol{x}) \tag{49}$$

$$\dot{\hat{\boldsymbol{x}}} = \boldsymbol{A}_0 \hat{\boldsymbol{x}} + \boldsymbol{f}_1(\hat{\boldsymbol{x}}) + \boldsymbol{K}_f(\hat{\boldsymbol{x}}) \left[\boldsymbol{y}(\boldsymbol{x}) - \boldsymbol{H}\hat{\boldsymbol{x}} \right] \tag{50}$$

where

$$\boldsymbol{f}_1(\boldsymbol{x}) = \boldsymbol{A}(\boldsymbol{x})\boldsymbol{x} \tag{51}$$

Define

$$\boldsymbol{e} = \hat{\boldsymbol{x}} - \boldsymbol{x} \tag{52}$$

Then

$$\begin{aligned} \dot{\boldsymbol{e}} &= \dot{\hat{\boldsymbol{x}}} - \dot{\boldsymbol{x}} = \boldsymbol{A}_0 \boldsymbol{e} + \left[\boldsymbol{f}_1(\hat{\boldsymbol{x}}) - \boldsymbol{f}_1(\boldsymbol{x}) \right] - \boldsymbol{K}_f(\hat{\boldsymbol{x}}) \boldsymbol{H} \boldsymbol{e} \\ &= \left[\boldsymbol{A}_0 - \boldsymbol{K}_f(\hat{\boldsymbol{x}}) \boldsymbol{H} \right] \boldsymbol{e} + \left[\boldsymbol{f}_1(\hat{\boldsymbol{x}}) - \boldsymbol{f}_1(\boldsymbol{x}) \right] \end{aligned}$$

Using (12) and (13) in $\dot{\boldsymbol{e}}$ yields

$$\begin{aligned} \dot{\boldsymbol{e}} &= \left[\boldsymbol{A}_0 - \sum_{i=0}^{\infty} \bar{\boldsymbol{T}}_i \boldsymbol{H}^T \boldsymbol{V}^{-1} \boldsymbol{H} \right] \boldsymbol{e} + \left[\boldsymbol{f}_1(\hat{\boldsymbol{x}}) - \boldsymbol{f}_1(\boldsymbol{x}) \right] \\ &= \left[\boldsymbol{A}_0 - \hat{\boldsymbol{T}}_0 \boldsymbol{H}^T \boldsymbol{V}^{-1} \boldsymbol{H} \right] \boldsymbol{e} \\ &\quad - \sum_{i=1}^{\infty} \bar{\boldsymbol{T}}_i \boldsymbol{H}^T \boldsymbol{V}^{-1} \boldsymbol{H} \boldsymbol{e} + \left[\boldsymbol{f}_1(\hat{\boldsymbol{x}}) - \boldsymbol{f}_1(\boldsymbol{x}) \right] \end{aligned} \tag{53}$$

Recall that $\sum_{i=0}^{\infty} \hat{\boldsymbol{T}}_i(\hat{\boldsymbol{x}}, \theta)\theta^i = \sum_{i=0}^{\infty} \bar{\boldsymbol{T}}_i(\hat{\boldsymbol{x}})$. From the condition (ii) in Theorem 3.1, we can always make $\boldsymbol{A}_0 - \hat{\boldsymbol{T}}_0 \boldsymbol{H}^T \boldsymbol{V}^{-1} \boldsymbol{H}$ a Hurwitz matrix. Let $\boldsymbol{F}_0 = \boldsymbol{A}_0 - \hat{\boldsymbol{T}}_0 \boldsymbol{H}^T \boldsymbol{V}^{-1} \boldsymbol{H}$.

Thus, for any given positive definite matrix $\hat{\boldsymbol{Q}} \in \mathrm{R}^{nxn}$, there exists a unique positive-definite $\hat{\boldsymbol{P}} \in \mathrm{R}^{nxn}$ such that

$$\boldsymbol{F}_0^T \hat{\boldsymbol{P}} + \hat{\boldsymbol{P}} \boldsymbol{F}_0 = -2\hat{\boldsymbol{Q}} \tag{54}$$

Now consider the following positive definite Lyapunov function

$$V(\boldsymbol{e}) = \boldsymbol{e}^T \hat{\boldsymbol{P}} \boldsymbol{e} \tag{55}$$

Then

$$\dot{V}(\boldsymbol{e}) = \dot{\boldsymbol{e}}^T \hat{\boldsymbol{P}} \boldsymbol{e} + \boldsymbol{e}^T \hat{\boldsymbol{P}} \dot{e} \tag{56}$$

$$\begin{aligned} \dot{V}(\boldsymbol{e}) = &\left\{ \left[\boldsymbol{e}^T \boldsymbol{F}_0^T - \boldsymbol{e}^T \boldsymbol{H}^T \boldsymbol{V}^{-1} \boldsymbol{H} \sum_{i=1}^{\infty} \bar{\boldsymbol{T}}_i \right. \right. \\ &\left. \left. + \left[\boldsymbol{f}_1^T(\hat{\boldsymbol{x}}) - \boldsymbol{f}_1^T(\boldsymbol{x}) \right] \right] \right\} \hat{\boldsymbol{P}} \boldsymbol{e} \\ &+ \boldsymbol{e}^T \hat{\boldsymbol{P}} \left\{ \left[\boldsymbol{F}_0 \boldsymbol{e} - \sum_{i=1}^{\infty} \bar{\boldsymbol{T}}_i \boldsymbol{H}^T \boldsymbol{V}^{-1} \boldsymbol{H} \boldsymbol{e} \right] + \left[\boldsymbol{f}_1(\hat{\boldsymbol{x}}) - \boldsymbol{f}_1(\boldsymbol{x}) \right] \right\} \end{aligned}$$

$$= -2\boldsymbol{e}^T\hat{\boldsymbol{Q}}\boldsymbol{e} + 2\boldsymbol{e}^T\hat{\boldsymbol{P}}\left\{\left[\boldsymbol{f}_1(\hat{\boldsymbol{x}}) - \boldsymbol{f}_1(\boldsymbol{x})\right] - \sum_{i=1}^{\infty}\bar{\boldsymbol{T}}_i\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\boldsymbol{e}\right\} \tag{57}$$

Since $\boldsymbol{f}_1(\boldsymbol{x})$ is Lipschitz continuous on a compact set Ω, it is easy to see that $\|\boldsymbol{f}_1(\hat{\boldsymbol{x}}) - \boldsymbol{f}_1(\boldsymbol{x})\| \le L_f\|\hat{\boldsymbol{x}} - \boldsymbol{x}\| = L_f\|\boldsymbol{e}\|$, where L_f is the Lipschitz constant. Then we have

$$\begin{aligned}\dot{V}(\boldsymbol{e}) &\le -2\lambda_{\min}(\hat{\boldsymbol{Q}})\cdot\|\boldsymbol{e}\|^2 + 2\cdot L_f\cdot\|\hat{\boldsymbol{P}}\|\cdot\|\boldsymbol{e}\|^2 \\ &\quad + 2\cdot\left\|\sum_{i=1}^{\infty}\bar{\boldsymbol{T}}_i\right\|\cdot\left\|\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\right\|\cdot\|\hat{\boldsymbol{P}}\|\cdot\|\boldsymbol{e}\|^2 \\ &= -2\left\{\lambda_{\min}(\hat{\boldsymbol{Q}}) - \|\hat{\boldsymbol{P}}\|\cdot\left[L_f + \left\|\sum_{i=1}^{\infty}\bar{\boldsymbol{T}}_i\right\|\cdot C_H\right]\right\}\|\boldsymbol{e}\|^2\end{aligned} \tag{58}$$

We have shown from (45) that $\bar{\boldsymbol{T}}_i$ satisfies

$$\|\bar{\boldsymbol{T}}_i\| \le (\varepsilon_1\cdots\varepsilon_i)C^iC_1^iC_2\cdots C_i\|\hat{\boldsymbol{T}}_0\| \quad (i = 1, 2, \cdots) \tag{59}$$

Therefore, as long as we choose proper $\varepsilon_1\cdots\varepsilon_i$ for $\|\sum_{i=1}^{\infty}\bar{\boldsymbol{T}}_i\|$ and large enough $\lambda_{\min}(\hat{\boldsymbol{Q}})$ such that

$$\|\hat{\boldsymbol{P}}\|\cdot\left[L_f + \left\|\sum_{i=1}^{\infty}\bar{\boldsymbol{T}}_i\right\|C_H\right] < \lambda_{\min}(\hat{\boldsymbol{Q}}), \tag{60}$$

then $\dot{V}(\boldsymbol{e}) < 0$.

Therefore, $\boldsymbol{e} = \boldsymbol{0}$ is an asymptotically stable equilibrium point. □

Due to the similarity between the θ-D observer and the SDREF, a comparison of these two is of interest. The SDREF formulation also brings the original nonlinear system (8) into the state dependent coefficient form [16]:

$$\dot{\boldsymbol{x}} = \boldsymbol{F}(\boldsymbol{x})\boldsymbol{x} \tag{61}$$

The SDREF is given by

$$\dot{\hat{\boldsymbol{x}}} = \boldsymbol{F}(\hat{\boldsymbol{x}})\hat{\boldsymbol{x}} + \boldsymbol{K}_f(\hat{\boldsymbol{x}})\left[\boldsymbol{y}(\boldsymbol{x}) - \boldsymbol{H}\hat{\boldsymbol{x}}\right] \tag{62}$$

where

$$\boldsymbol{K}_f(\hat{\boldsymbol{x}}) = \hat{\boldsymbol{P}}_s(\hat{\boldsymbol{x}})\boldsymbol{H}^T\boldsymbol{V}^{-1} \tag{63}$$

and $\hat{\boldsymbol{P}}_s(\hat{\boldsymbol{x}})$ is the positive definite solution to the state dependent Riccati equation:

$$\boldsymbol{F}(\hat{\boldsymbol{x}})\hat{\boldsymbol{P}}_s(\hat{\boldsymbol{x}}) + \hat{\boldsymbol{P}}_s(\hat{\boldsymbol{x}})\boldsymbol{F}^T(\hat{\boldsymbol{x}}) - \hat{\boldsymbol{P}}_s(\hat{\boldsymbol{x}})\boldsymbol{H}^T\boldsymbol{V}^{-1}\boldsymbol{H}\hat{\boldsymbol{P}}_s(\hat{\boldsymbol{x}}) + \boldsymbol{W} = \boldsymbol{0} \tag{64}$$

Comparing Eqs. (10)–(13) with Eqs. (61)–(64), one can see that the θ-D observer takes a form similar to the SDREF. The solution $\sum_{i=0}^{\infty}\hat{\boldsymbol{T}}_i(\hat{\boldsymbol{x}},\theta)\theta^i$ of the θ-D observer is equivalent to the solution of the state dependent Riccati equation, $\hat{\boldsymbol{P}}_s(\hat{\boldsymbol{x}})$. The advantage of the θ-D observer over the SDREF is the analytical computation of the observer gain (13). The SDREF gain though, needs solution of the algebraic Riccati equation (64) at every instant, which is computationally intensive.

It is also noted that the θ-D observer is different from the continuous EKF in the form of $\dot{\hat{\boldsymbol{x}}} = \boldsymbol{F}(\hat{\boldsymbol{x}})\hat{\boldsymbol{x}} + \boldsymbol{K}_f(\hat{\boldsymbol{x}})[\boldsymbol{y}(\boldsymbol{x}) - \boldsymbol{H}\hat{\boldsymbol{x}}]$ because $\boldsymbol{F}(\hat{\boldsymbol{x}})$ is not obtained from linearization about the current estimate. The θ-D observer addresses the nonlinear dynamics directly without approximation. It also demands much less computations than EKF since EKF requires calculation of the Jacobian matrix and either to integrate the nonlinear covariance matrix or to solve the filter Riccati equation continuously to calculate the gains whereas the θ-D observer gains are computed analytically.

4. Unified flight control and estimation of high performance aircraft

4.1. Unified optimal control and observer design algorithm

The unified optimal control and observer designs described in the last two sections can be summarized in the following algorithm, which is given as the block diagram in Fig. 1.

This algorithm can be applied to general control-affine nonlinear systems with linear measurement. As can be seen, the optimal control and observer designs are unified by the same θ-D algorithm. When the system matrices $\boldsymbol{A}_0, \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{Q}, \boldsymbol{R}$ are the inputs to the θ-D algorithm, closed-form optimal control gain can be obtained. When the dual of the system matrices $\boldsymbol{A}_0, \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{Q}, \boldsymbol{R}$, i.e. $\boldsymbol{A}_0^T, \boldsymbol{A}^T, \boldsymbol{H}^T, \boldsymbol{W}, \boldsymbol{V}$ are the input to the θ-D algorithm, closed-form observer gain can be obtained.

Remark 4.1. In this unified θ-D optimal control and observer design, the separation principle is implicitly assumed, which usually does not hold for general nonlinear systems. Therefore, optimality of this design scheme cannot be claimed in this paper and the theoretical stability of the closed-loop system with the observer in the feedback will be investigated in the future work.

4.2. Flight control of high performance aircraft

In this section, the θ-D observer and control are applied to the flight control of a high performance aircraft operating at high angles of attack. The mathematical model used in this study is similar to the X-31 research aircraft [5] and only the longitudinal mode is considered. The longitudinal model includes nonlinear aerodynamic stability and control derivatives that are functions of angle of attack. Thus, at high angle of attack, the design model is highly nonlinear and the conventional linear time-invariant framework is not applicable. The state vector that describes the longitudinal motion is

$$\boldsymbol{x} = [\,\Delta V \quad \Delta\alpha \quad q \quad \gamma \quad \Delta\delta\,]^T \tag{65}$$

where ΔV is the deviation of the velocity from the level flight trim value of 100 m/s; $\Delta\alpha$ is the deviation of angle of attack from its trim value of 4.2°; q is the pitch rate in rad/s; γ is the flight path angle in radians; $\Delta\delta$ is the change in canard deflection in degrees from its trim value. The scalar control u is the input to the canard actuator. The canard is the aerodynamic control surface that can deflect the nose down by as much as 90° to remain unstalled at high wing angles of attack. The aircraft considered in this study is so unstable as to be unflyable without feedback control systems. The longitudinal equations of motion are given by

$$\dot{\boldsymbol{x}} = (\boldsymbol{A}_L + x_2\boldsymbol{A}_{NL})\boldsymbol{x} + \boldsymbol{B}u \tag{66}$$

where the matrices $\boldsymbol{A}_L$ and $\boldsymbol{A}_{NL}$ were obtained by a best least-squares fit to flight conditions at eight angles of attack ranging from 4.2° to 43° and are given below [5]:

$$\boldsymbol{A}_L = \begin{bmatrix} -0.0443 & 112.80 & 0.0 & -9.807 & 0.0 \\ -0.00049 & -2.5390 & 1.0 & 0.0 & -0.00149 \\ -0.00073 & 19.3200 & -2.2700 & 0.0 & 0.39590 \\ 0.00049 & 2.53900 & 0.0 & 0.0 & 0.00149 \\ 0.0 & 0.0 & 0.0 & 0.0 & 20.0 \end{bmatrix}$$

$$\boldsymbol{A}_{NL} = \begin{bmatrix} -0.23171 & -0.00109 & 0.0 & 0.0 & 0.0 \\ -0.012760 & -0.79219 & 0.0 & 0.0 & 0.00036 \\ 0.00102 & 64.2940 & -13.9710 & 0.0 & -0.09454 \\ 0.012760 & 0.79219 & 0.0 & 0.0 & -0.00036 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}$$

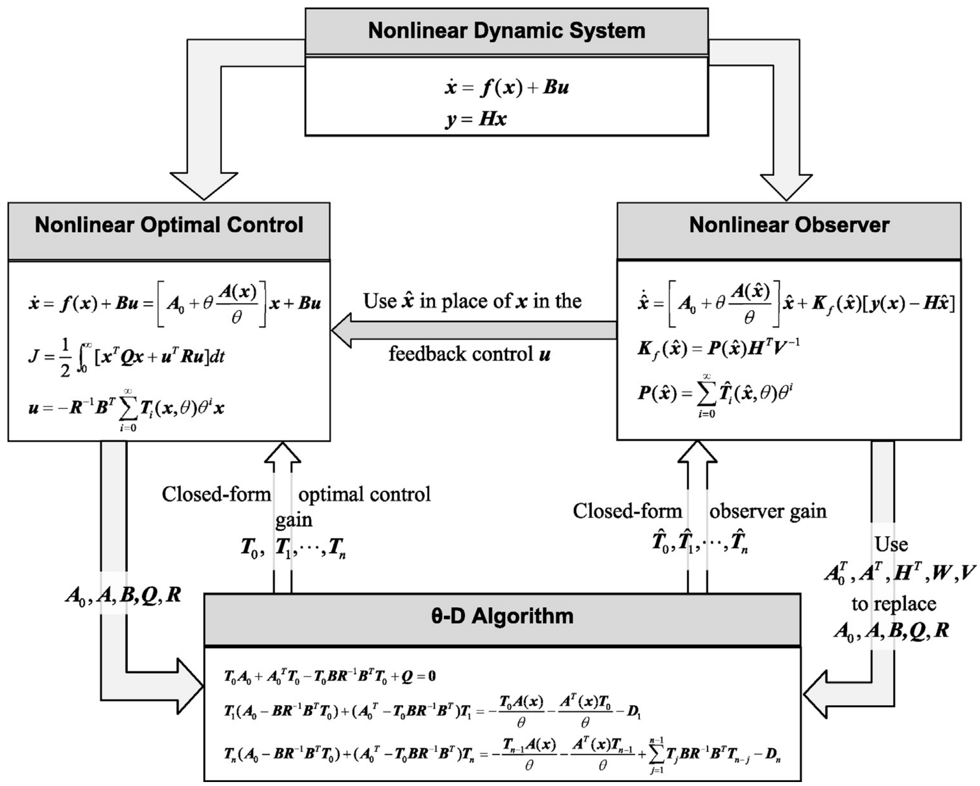

Fig. 1. Unified θ-D optimal control and observer design algorithm.

$$\boldsymbol{B} = [\,0.0 \quad 0.0 \quad 0.0 \quad 0.0 \quad 20.0\,]^T.$$

The objective of this study is to design an optimal feedback controller and the observer to control the angle of attack and flight path angle.

The available measurements are assumed to be the velocity and the canard deflection, i.e.

$$\boldsymbol{y} = [\,x_1 \quad x_5\,]^T \tag{67}$$

To employ the θ-D technique, the nonlinear equations need to be written in a linear-like structure (3),

$$\dot{\boldsymbol{x}} = \boldsymbol{F}(\boldsymbol{x})\boldsymbol{x} + \boldsymbol{B}u(\hat{\boldsymbol{x}}) = \left[\boldsymbol{A}_0 + \theta\frac{\boldsymbol{A}(\boldsymbol{x})}{\theta}\right]\boldsymbol{x} + \boldsymbol{B}u(\hat{\boldsymbol{x}}) \tag{68}$$

where

$$\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{A}_L + x_2\boldsymbol{A}_{NL}$$

Note that the feedback state $\hat{\boldsymbol{x}}$ in the control will be the outputs of the θ-D observer.

$\boldsymbol{A}_0$ and $\boldsymbol{A}(\boldsymbol{x})$ in Eq. (68) are chosen in the following way:

$$\boldsymbol{A}_0 = \boldsymbol{F}\big(\boldsymbol{x}(t_0)\big) \quad \text{and} \quad \boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{F}(\boldsymbol{x}) - \boldsymbol{F}\big(\boldsymbol{x}(t_0)\big)$$

The cost function is in a quadratic form of (2). The optimal control is designed using the θ-D technique

$$u(\hat{\boldsymbol{x}}) = -\boldsymbol{R}^{-1}\boldsymbol{B}^T\big[\boldsymbol{T}_0 + \boldsymbol{T}_1(\hat{\boldsymbol{x}},\theta)\theta + \boldsymbol{T}_2(\hat{\boldsymbol{x}},\theta)\theta^2\big]\hat{\boldsymbol{x}} \tag{69}$$

where $\boldsymbol{T}_i(\hat{\boldsymbol{x}},\theta)$ is obtained by following the algorithm (5) and the feedback state vector $\hat{\boldsymbol{x}}$ is estimated using the θ-D observer

$$\dot{\hat{\boldsymbol{x}}} = \boldsymbol{F}(\hat{\boldsymbol{x}})\hat{\boldsymbol{x}} + \boldsymbol{B}u(\hat{\boldsymbol{x}}) + \boldsymbol{K}_f(\hat{\boldsymbol{x}})\big[\boldsymbol{y}(\boldsymbol{x}) - \boldsymbol{H}\hat{\boldsymbol{x}}\big] \tag{70}$$

The observer gain $\boldsymbol{K}_f(\hat{\boldsymbol{x}})$ is calculated by the observer algorithm (12)–(15). In this problem, the first three terms in the θ-D control (69) and observer algorithms (13) were found sufficient to yield good performance. Note that both the optimal control law $u(\hat{\boldsymbol{x}})$ and the observer gain $\boldsymbol{K}_f(\hat{\boldsymbol{x}})$ can be solved in a closed-form that offers a great computational advantage.

The first simulation scenario is to regulate the states to zero (their respective trim values) from a large initial angle of attack given by $\boldsymbol{x}_0 = [\,0\ 30°\ 0\ 0\ 0\,]^T$. The initial estimated states are assumed to be: $\hat{\boldsymbol{x}}(0) = [\,0\ 25°\ 0\ 0\ 0\,]^T$.

After some numerical experiments, the control weights and the observer weights are chosen to be:

$$\boldsymbol{Q} = \text{diag}(80, 1, 1, 300, 1), \quad R = 300$$
$$\boldsymbol{W} = I_5, \quad \boldsymbol{V} = \text{diag}(0.1, 0.1) \tag{71}$$

where diag(.) represents a diagonal matrix and I_5 represents an identity matrix of dimension 5. The k_i and l_i parameters in the perturbation matrices $D_1(\hat{D}_1)$ and $D_2(\hat{D}_2)$ for both control and observer are chosen to be $k_1 = k_2 = 1$, and $l_1 = l_2 = 0.01$.

Figs. 2–5 show the time responses of the velocity, angle of attack, pitch rate, and flight path angle respectively. The solid line shows the actual state trajectory employing the θ-D optimal control with the feedback states estimated by the θ-D observer. The dashed line shows the estimated state trajectories. As can be seen, the estimated states converge to the actual states very quickly. The optimal control regulates the states to zero with good transient responses. The actual velocity and the estimated velocity are too close to be distinguishable in Fig. 1 because the velocity is a state that can be directly measured.

For comparison, the extended Kalman filter (EKF) is used as a nonlinear observer for the θ-D optimal control. Results with the same controller design are shown in Figs. 6 and 7 where only angle of attack and flight path angle responses are presented for brevity. As can be seen in Fig. 6, the angle of attack estimate also converges to the actual one quickly. Comparing Fig. 7 to Fig. 5, the flight path angle estimate using the EKF has a better initial transient response but converges more slowly to the actual as compared to the θ-D

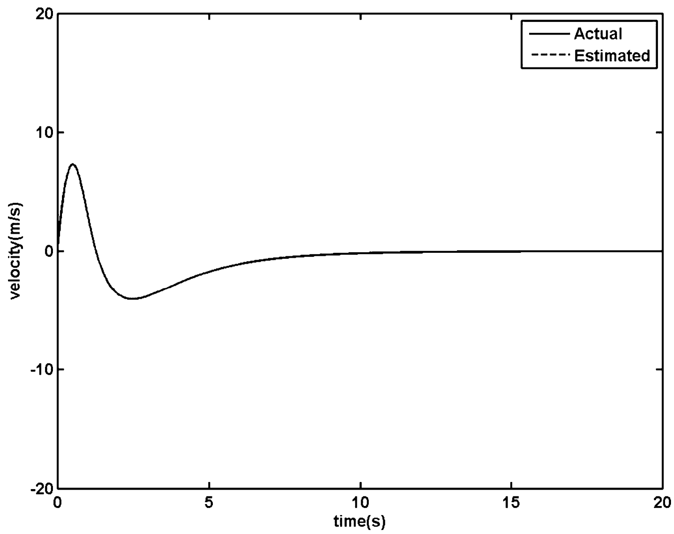

Fig. 2. Velocity regulation using the θ-D observer.

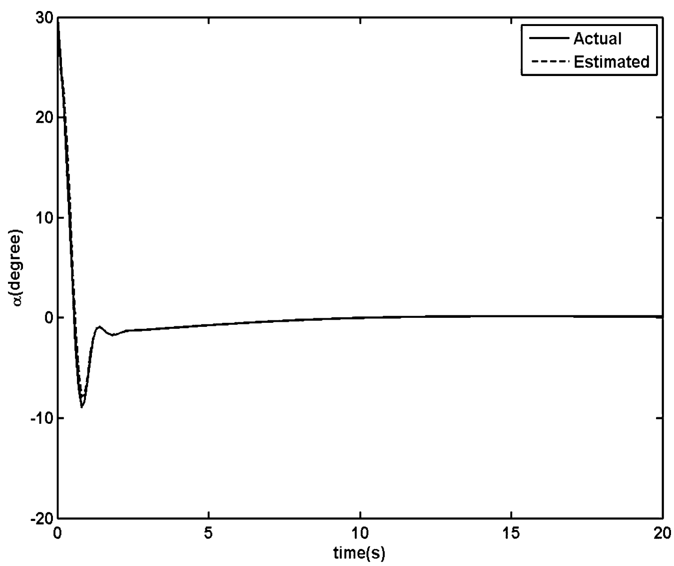

Fig. 3. Angle of attack regulation using the θ-D observer.

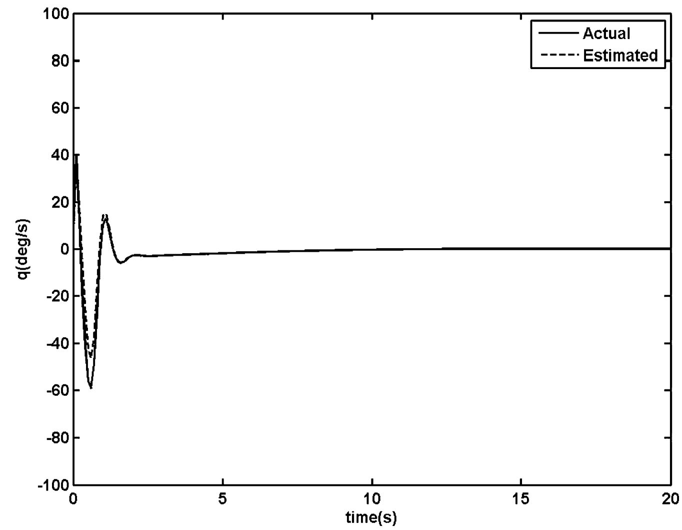

Fig. 4. Pitch rate regulation using the θ-D observer.

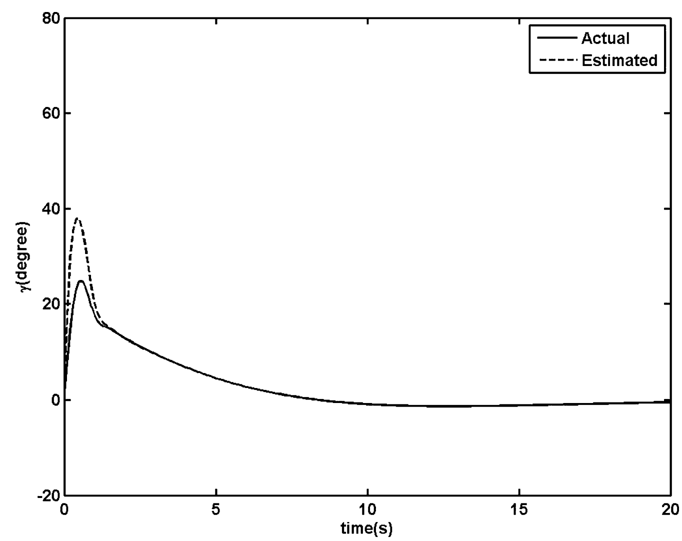

Fig. 5. Flight path angle regulation using the θ-D observer.

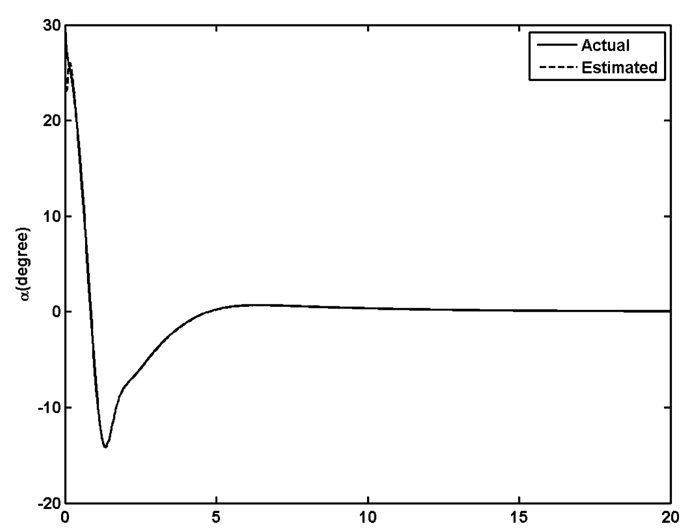

Fig. 6. Angle of attack regulation using EKF.

observer performance. Also, both angle of attack and flight path angle using the EKF are regulated to zero relatively more slowly than using the θ-D observer. Fig. 8 gives the canard deflection and control command histories (i.e. optimal control) using the θ-D observer and the EKF. Both of them show satisfactory responses. Overall, the EKF also produces comparably good results for this scenario.

The second scenario is to track a command flight path angle γ_c. The commanded γ_c starts from 0 until 5 seconds and gradually increases to 45°, holds until 15 seconds, and then returns to 0 at 20 seconds as shown in Fig. 9. In order to ensure a good tracking performance, an integral state of the flight path angle is augmented into the original state, i.e.

$$\dot{\gamma}_I = \gamma \tag{72}$$

The state space for this tracking problem becomes

$$\boldsymbol{x} = [\,\Delta V \quad \Delta\alpha \quad q \quad \gamma \quad \Delta\delta \quad \gamma_I\,]^T \tag{73}$$

and the associated $\boldsymbol{A}_L$ and $\boldsymbol{A}_{NL}$ matrices are changed accordingly.

The optimal control is applied as a servomechanism for this tracking problem [26], i.e.

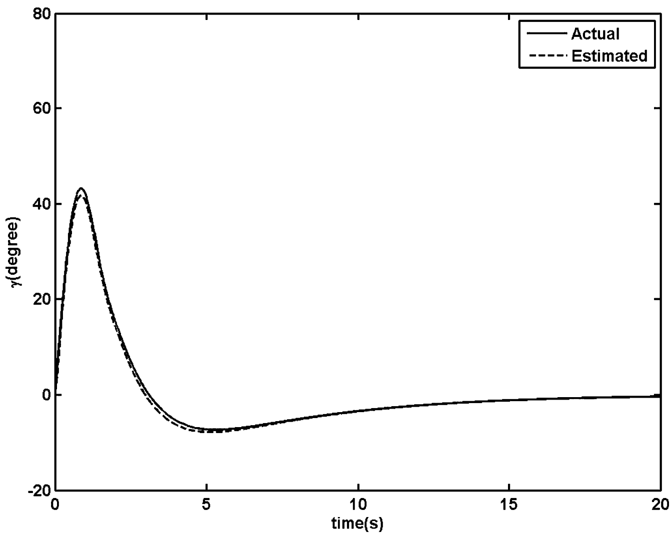

Fig. 7. Flight path angle regulation using EKF.

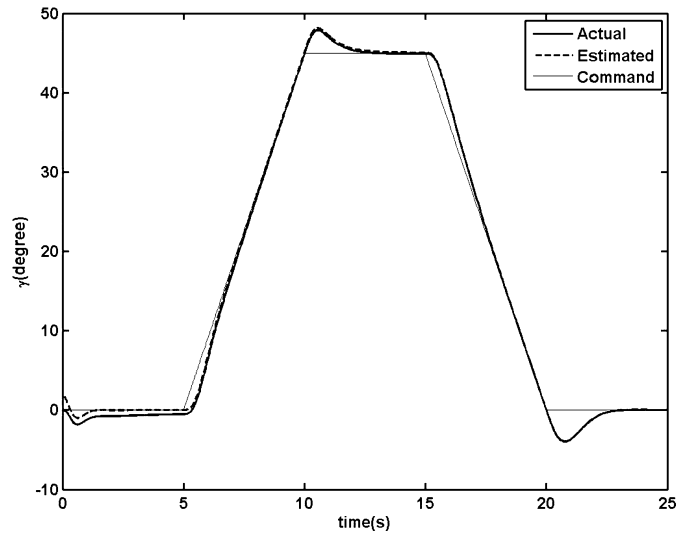

Fig. 9. Flight path angle tracking using the θ-D observer.

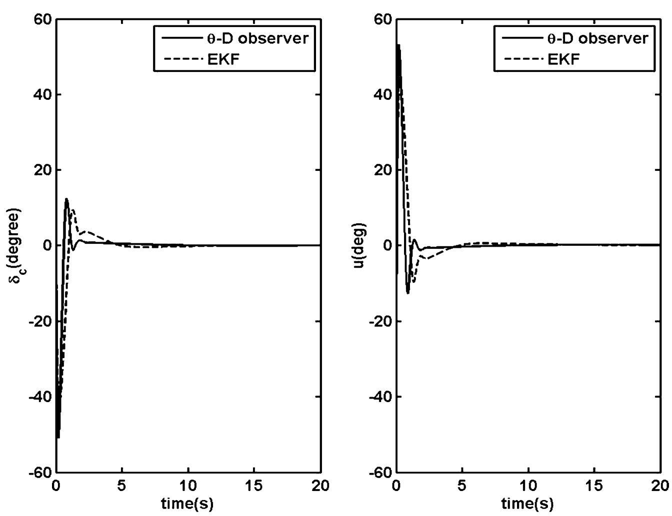

Fig. 8. Canard deflection and control command histories using θ-D observer and EKF.

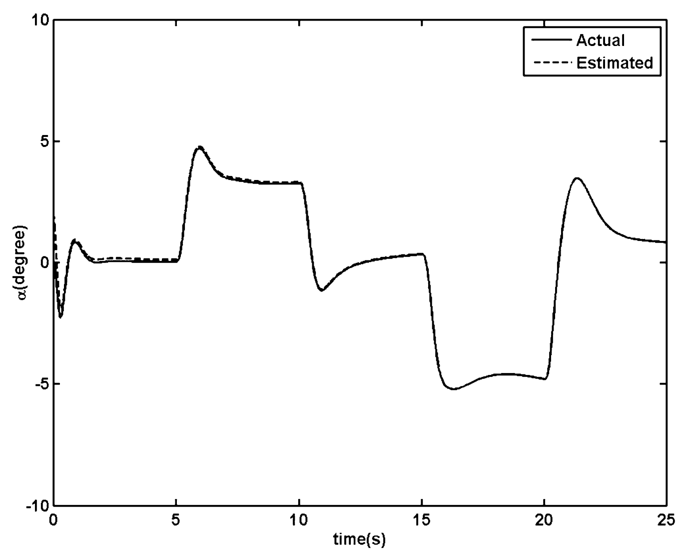

Fig. 10. Angle of attack tracking response using the θ-D observer.

$$u = -\boldsymbol{R}^{-1}\boldsymbol{B}^{T}\left[\boldsymbol{T}_0 + \boldsymbol{T}_1(\hat{\boldsymbol{x}}, \theta)\theta + \boldsymbol{T}_2(\hat{\boldsymbol{x}}, \theta)\theta^2\right][\hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}_r]^T \tag{74}$$

where $\hat{\boldsymbol{x}}_r = [0\ 0\ 0\ \gamma_c\ 0\ \int \gamma_c\,]^T$ is the reference state vector.

Initial states are set at zero and the initial estimated states are set slightly off the actual, i.e. $\hat{\boldsymbol{x}}(0) = [0\ 2°\ 0\ 2°\ 0\,]^T$. Note that the observer state space is the same as in the first scenario because the integral state does not need to be estimated.

Weights used for control and estimation are chosen to be

$$\boldsymbol{Q} = \mathrm{diag}(1, 1, 1, 100, 1, 200), \quad R = 10,$$

$$\boldsymbol{W} = I_5, \quad \boldsymbol{V} = \mathrm{diag}(0.01, 0.01).$$

The k_i and l_i parameters are the same as those in the first scenario.

Figs. 9 and 10 demonstrate the flight path angle response and the angle of attack response respectively, which show that the flight path angle tracks the command very well and both estimated angles from the θ-D observer converge to the actual values very quickly.

For comparison, the EKF is also used for state estimation in this flight path angle tracking scenario. The difference between these two observers is significant as shown in Figs. 11 and 12. The flight path angle tracking and the angle of attack response

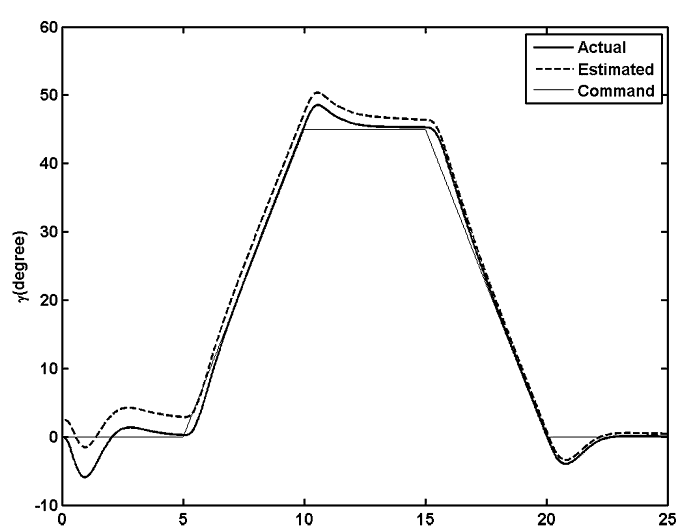

Fig. 11. Flight path angle tracking using EKF.

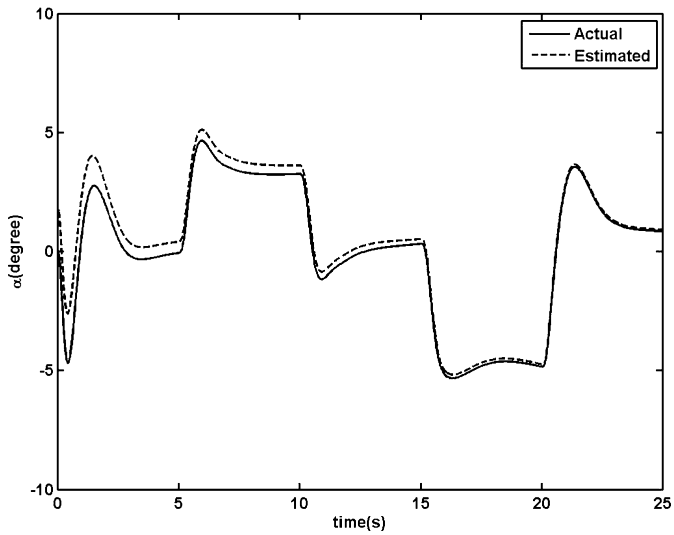

Fig. 12. Angle of attack tracking using EKF.

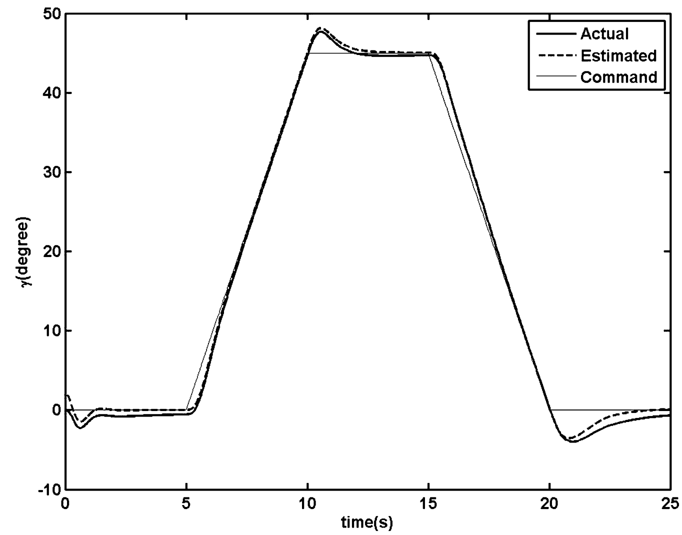

Fig. 14. Flight path angle tracking under +15% modeling uncertainty.

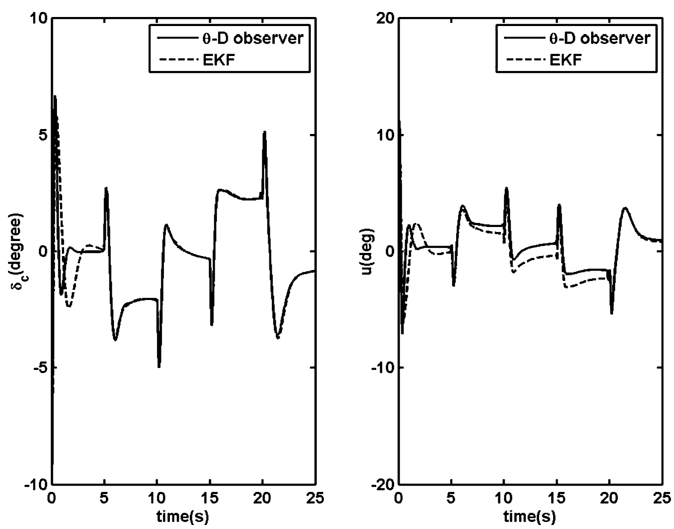

Fig. 13. Canard deflection and control command histories using θ-D observer and EKF.

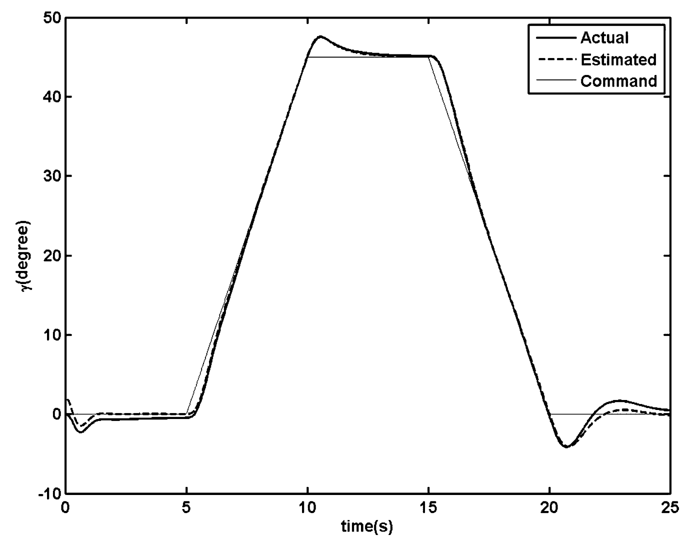

Fig. 15. Flight path angle tracking under −15% modeling uncertainty.

are obviously worse. Also, the estimates using the EKF converge to the actual states considerably more slowly than those using the θ-D observer. Typically, these can happen due to the linearization involved in the EKF process. Better performance of the θ-D observer can be attributed to the fact that its formulation does not involve a linearization process in spite of a state-dependent linear like structure. Besides the tracking accuracy, there is a remarkable difference between the EKF and the θ-D observer in terms of computations. For EKF, either the nonlinear covariance matrix needs to be numerically integrated or the filtering Riccati equation has to be solved continuously to calculate the gain. The computation for either option is numerically intensive and grows significantly with the size of the state space. In contrast, however, the θ-D gains are computed analytically. The canard deflection and control command histories using the θ-D observer and EKF are shown in Fig. 13. Both exhibit reasonable response.

In order to test the robustness of the θ-D method based flight control and observer, the aircraft model is assumed to have uncertainty. Specifically, the nonlinear coefficient matrix $\boldsymbol{A}_{NL}$ is assumed to have $\pm 15\%$ uncertainty. The flight control and observer are still designed based on the nominal model. Figs. 14 and 15 show the flight path angle tracking response under this uncertainty. Compared to the normal case in Fig. 9, only slight degrading of tracking performance can be observed, mainly in the last five seconds, and the θ-D method shows satisfactory robustness.

5. Conclusions

In this paper, a unified θ-D formulation was utilized to design both the nonlinear optimal controller and the observer based on the duality concept. The major advantage of this approach is that the optimal control and observer gains can be obtained as closed-form expressions using the same θ-D algorithm and consequently do not need complex on-line computations. The θ-D observer does not involve a linearization process required by EKF and addresses the nonlinearity directly. The unified θ-D optimal control and observer are applied to the flight control of a highly maneuverable aircraft with high nonlinearity and unstable dynamics. The effectiveness of this technique has been demonstrated by regulating the high angle of attack and tracking flight path angle command. The simulation also shows the favorable results compared with the widely used extended Kalman filter.

Conflict of interest statement

None declared.

Acknowledgement

This work was supported by the U.S. National Science Foundation CAREER Award ECCS-0846877.

References

[1] S.C. Beeler, H.T. Tran, H.T. Banks, Feedback control methodologies for nonlinear systems, J. Optim. Theory Appl. 107 (1) (2000) 1–33.
[2] A.E. Bryson, Y.-C. Ho, Applied Optimal Control, Hemisphere Publishing Corporation, 1975.
[3] T. Cimen, Survey of state-dependent Riccati equation in nonlinear optimal feedback control synthesis, J. Guid. Control Dyn. 35 (4) (2012) 1025–1947.
[4] M.A. Dornheim, X-31, F-16 MATV, F/A-18 HARV explore diverse missions, Aviat. Week Space Technol. 140 (1994) 46–47.
[5] W.L. Garrard, D.F. Enns, S.A. Snell, Nonlinear feedback control of highly maneuverable aircraft, Int. J. Control 56 (4) (1992) 799–812.
[6] A. Gelb, Applied Optimal Estimation, M.I.T. Press, Cambridge, MA, 1974.
[7] A. Glumineau, C. Moog, F. Plestan, New algebraic-geometric conditions for the linearization by input–output injection, IEEE Trans. Autom. Control 41 (4) (1996) 598–603.
[8] R.R. Harman, I.Y. Bar-Itzhack, Pseudolinear and state-dependent Riccati equation filters for angular rate Estimation, J. Guid. Control Dyn. 22 (5) (1999) 723–725.
[9] W.B. Herbst, Future fighter technologies, J. Aircr. 17 (8) (1980) 561–566.
[10] M. Hou, A. Pugh, Observer with linear error dynamics for nonlinear multi-input systems, Syst. Control Lett. 37 (1) (1999) 1–9.
[11] A. Isidori, Nonlinear Control Systems: An Introduction, 3rd edition, Springer-Verlag, London, 1995.
[12] B.S. Kim, A.J. Calise, Nonlinear flight control using neural networks, J. Guid. Control Dyn. 20 (1) (1995) 26–33.
[13] A.F. Lynch, S.A. Bortoff, Nonlinear observers with approximately linear error dynamics: the multivariable case, IEEE Trans. Autom. Control 46 (6) (2001) 927–932.
[14] T. Mori, I.A. Derese, A brief summary of the bounds on the solution of the algebraic matrix equations in control theory, Int. J. Control 39 (2) (1984) 247–256.
[15] C.P. Mracek, J.R. Cloutier, Control designs for the nonlinear benchmark problem via the State-Dependent Riccati Equation Method, Int. J. Robust Nonlinear Control 8 (1998) 401–433.
[16] C.P. Mracek, J.R. Cloutier, C.N. D'Souza, A new technique for nonlinear estimation, in: Proceedings of the IEEE Conference on Control Applications, Dearborn, MI, September 1996.
[17] D. Noh, N.H. Jo, J.H. Seo, Nonlinear observer design by dynamic observer error linearization, IEEE Trans. Autom. Control 49 (10) (2004) 1750–1796.
[18] R. Rajamani, Observers for Lipschitz nonlinear systems, IEEE Trans. Autom. Control 43 (3) (1998) 397–401.
[19] J. Reiner, G.J. Balas, W.L. Garrard, Robust dynamic inversion for control of highly maneuverable aircraft, J. Guid. Control Dyn. 18 (1) (1995) 18–24.
[20] K. Robenack, A.F. Lynch, High-gain nonlinear observer design using the observer canonical form, IET Control Theory Appl. 1 (6) (2007) 1574–1579.
[21] H. Shim, J.H. Seo, Recursive nonlinear observer design: beyond the uniform observability, IEEE Trans. Autom. Control 48 (2) (2003) 294–298.
[22] Y. Shin, A.J. Calise, M.D. Johnson, Adaptive control of advanced fighter aircraft in nonlinear flight regimes, J. Guid. Control Dyn. 31 (5) (2008) 1464–1477.
[23] F.E. Thau, Observing the state of nonlinear dynamical systems, Int. J. Control 17 (3) (1973) 471–479.
[24] Q. Wang, R.F. Stengel, Robust nonlinear flight control of a high-performance aircraft, IEEE Trans. Control Syst. Technol. 13 (1) (2005) 15–26.
[25] M. Xin, S.N. Balakrishnan, A new method for suboptimal control of a class of nonlinear systems, Optim. Control Appl. Methods 26 (2) (2005) 55–83.
[26] M. Xin, S.N. Balakrishnan, D.T. Stansbery, E.J. Ohlmeyer, Nonlinear missile autopilot design with θ-D technique, AIAA J. Guid. Control Dyn. 27 (3) (2004) 406–417.
[27] M. Xin, H. Pan, Indirect robust control of spacecraft via optimal control solution, IEEE Trans. Aerosp. Electron. Syst. 48 (2) (2012) 1798–1809.
[28] F. Zhu, Z. Han, A note on observers for Lipschitz nonlinear systems, IEEE Trans. Autom. Control 47 (10) (2002) 1751–1754.