Extended Kalman Filter-Based Observer Design for Semilinear Infinite-Dimensional Systems

Sepideh Afshar, Fabian Germ, and Kirsten Morris, Fellow, IEEE

Abstract—In many physical applications, the system’s state varies with spatial variables as well as time. The state of such systems is modeled by partial differential equations and evolves on an infinite-dimensional space. Systems modeled by delay-differential equations are also infinite-dimensional systems. The full state of these systems cannot be measured. Observer design is an important tool for estimating the state from available measurements. For linear systems, both finite- and infinite-dimensional, the Kalman filter provides an estimate with minimum-variance on the error, if certain assumptions on the noise are satisfied. The extended Kalman filter (EKF) is one type of extension to nonlinear finite-dimensional systems. In this article we provide an extension of the EKF to semilinear infinite-dimensional systems. Under mild assumptions we prove the well-posedness of equations defining the EKF. Next, local exponential stability of the error dynamics is shown. Only detectability is assumed, not observability, so this result is new even for finite-dimensional systems. The results are illustrated with implementation of finite-dimensional approximations of the infinite-dimensional EKF on an example.

Index Terms—Exponential stability, extended Kalman filter (EKF), infinite-dimensional systems, semilinear systems.

I. INTRODUCTION

In many physical applications, the system’s state varies with spatial variables as well as time. The state of such systems is modeled by partial differential equations and evolves on an infinite-dimensional space, and so they are an important class of infinite-dimensional systems. Systems modeled by delay-differential equations are also infinite-dimensional systems. The full state of these systems cannot be measured. As for finite-dimensional systems, a system, referred to as an observer or estimator, can be designed to estimate the state using the mathematical model and the measurements provided by sensors.

For linear systems, the Kalman filter (KF) minimizes the variance of the error under certain assumptions on the disturbances. The observer can be calculated through solution of a Riccati equation. The KF is widely used and was extended to infinite-dimensional linear systems in the 1970s; see [1], [2], and [3], and also the book [4]. This theory was recently extended to time-varying infinite-dimensional systems [5]. For linear infinite-dimensional systems, there are a number of other different approaches to observer design in addition to the KF, including backstepping [6], sliding mode combined with backstepping [7]. Some other approaches can be found in [8], [9], [10], and [11].

Due to its success in a wide range of applications, an extension of the KF to nonlinear systems, the extended Kalman filter (EKF), was developed for finite-dimensional systems. The EKF design is based on a linear approximation of the system around the estimated state. The linearized system is used to calculate a KF that is updated as the system state evolves, e.g., [12], [13]. This method is widely used; see, for example, [14], [15], [16], [17], [18], and [19]. However, although this method may work well, it is well known that it may lead to divergent error estimates.

Multiple studies have been dedicated to convergence analysis of the state estimate by the EKF and its required condition for finite dimensional systems. In [20], the global asymptotic convergence of a discrete time linear difference system has shown under a uniform observability assumption. Similarly, local asymptotic convergence of the estimation error under a uniform observability assumption of the linearized system has been shown in [21] for discrete time systems. In [22], conditions for asymptotic convergence are imposed on the linearization residues as well as by a reconstructibility assumption, resembling an uniform observability assumption of the linearized system. Under an uniform observability condition it is shown in [15] that the estimation error is bounded in presence of disturbances for discrete and later for continuous time system [16]. Furthermore, local exponential convergence of the estimation error under an uniform controllability condition as well as a special type of uniform detectability assumption on the linearized system is shown in [23] for continuous time systems. Local exponential convergence of the estimation error with prescribed convergence rate is shown in [24] for continuous time and in [14] for discrete time under certain assumptions that imply observability. The stability of a EKF observer for a class of systems represented in a canonical form is also studied in [25]; here, the system is assumed to be globally asymptotically stable in addition with uniformly bounded linearized operator. Under these strong assumptions, it is shown that the EKF is globally exponentially stable. In general, in addition to the assumptions imposed on the linearized portion of the system, convergence of
the estimation error depends on the size of the nonlinearity and the initial condition, see for instance, [26] and [27].

Observers for nonlinear infinite-dimensional systems are often designed using a finite-dimensional approximation of the system. Some examples are the robust fuzzy and also robust adaptive observers in [28] and [29]. In [30], the effect of approximation on observer performance for several different types of diffusion models and different observer designs was studied. The EKF has been used on finite-dimensional approximations of PDEs; for example a highway traffic model in [31] and state-of-charge estimation in lithium-ion batteries [32].

There are some studies for nonlinear infinite-dimensional systems where the observer is designed directly using the infinite-dimensional system equations. In [33], a second-order sliding mode observer is employed to provide stability with the assumption that the measurement is available everywhere. In [34], the observer dynamics are corrected by a linear output error injection term via an estimated spatially distributed measurement. Spatially-distributed linear output injection is also proposed in [35] for a 1-D nonlinear Burgers’ equation. Backstepping is used in [10] to design an observer for a lithium-ion battery model using the PDE model directly. An example of a general and abstract form of late lumping nonlinear observer design is introduced in [36] on reflexive Banach spaces, where a nonlinear feedback operator is added to a copy of the system’s dynamics. Other examples of observer design for specific nonlinear PDEs can be found in [37], [38], [39], and [40].

However, there are no theoretical results for the EKF for infinite-dimensional systems. As for a finite-dimensional EKF, the observer dynamics are a copy of the original system’s dynamics with an injection gain defined by the solution of an Riccati equation. Since the Riccati equation is coupled with the observer equation, conventional results in the literature including [41] for existence of solutions to the Riccati equation cannot be directly used. This is due to the fact that for linear equations the Riccati equation does not depend on the state of the system. In our, nonlinear, case, such a dependence still remains after linearizing the system, making the analysis more involved. The rest of this article is organized as follows. After formally defining the problem and the class of systems to be considered in Section II, in Section III it is shown that the equations defining the EKF are well-posed and possess a unique solution for a class of semilinear infinite-dimensional systems with bounded observation. The proof of well-posedness was previously reported in [42] for nonlinearities without time dependence, and briefly sketched in the conference paper [43]. A complete proof with a slightly different presentation and considering time-dependent nonlinearities is provided in this article, and a more complex example is presented than in [43]. Section IV is devoted to some technical results involving stability of the time-varying problem.

In Section V the exponential convergence of the observer state to true state is shown. For sufficiently small initial error, and smooth nonlinearity, the error dynamics are proved to be exponentially stable. We also show that with these assumption, the estimation error is bounded in presence of disturbances. Our contribution in proving local exponential convergence of the error dynamics for EKF in the infinite-dimensional setting is twofold. First, finite-dimensional results depend on observability assumptions, the analogue of which would be uniform exact observability. Since exact observability is a very strong assumption for infinite-dimensional systems, the existing finite-dimensional proofs could not simply be extended to infinite-dimensional systems. This led us to obtain local exponential of the error dynamics under assumptions of uniform stabilizability/detectability which are weaker than those used in the existing finite-dimensional literature. Thus, these results are new even for finite-dimensional systems. Section VI shows an example in which a simple magnetic drug delivery system is considered. Finally, Section VII concludes this article with the illustration of implementation of this approach for estimation of concentration in a magnetic drug delivery system.

Notation: Throughout this article, calligraphic $\mathcal{H}$, with or without indices, will denote Hilbert spaces, where the space considered is clear from the context, the norm $\| \cdot \|$, as well as the scalar product $(\cdot, \cdot)$, will be equipped with an appropriate subscript, i.e., $\| \cdot \|_{\mathcal{H}}$ or $(\cdot, \cdot)_{\mathcal{H}}$; otherwise, it is omitted. For an arbitrary, but henceforth fixed $t_f > 0$ we work on the time interval $[0, t_f]$. For $0 < p < \infty$ let $L^p([0, t_f], \mathcal{H})$ and $W^{1,p}([0, t_f], \mathcal{H})$ denote the spaces of functions $f : [0, t_f] \to \mathcal{H}$ such that

$$
\int_0^{t_f} \| f(t) \|_{\mathcal{H}}^p \, ds < \infty \quad \text{and} \quad \int_0^{t_f} \| f(t) \|_{\mathcal{H}}^p + \| \frac{d}{dt} f(t) \|_{\mathcal{H}}^p \, ds < \infty
$$

respectively. By $\mathcal{H}_2 \to \mathcal{H}_1$ we denote a continuous and dense embedding and the trace space of $L^p([0, t_f]; \mathcal{H}_2) \cap W^{1,p}([0, t_f]; \mathcal{H}_1)$ is denoted by $\mathcal{H}_{2,1/p}$. For integers $k \geq 0$ we denote by $C^k([0, t_f], \mathcal{H}_2)$ the space of functions $f : [0, t_f] \to \mathcal{H}$ that are $k$ times continuously (with respect to the norm in $\mathcal{H}$) differentiable. Whenever $\mathcal{H} = \mathbb{R}$ we may omit the space and simply write $L([0, t_f]), W^{1,p}([0, t_f])$ and $C^k([0, t_f])$ for the spaces above. For a linear operator $A : \mathcal{H}_1 \to \mathcal{H}_2$ we write $D(A) \subset \mathcal{H}_1$ for its domain and denote by

$$
\| A \| = \sup_{v \in \mathcal{H}_2, \| v \| \leq 1} \| Av \|_{\mathcal{H}_2}
$$

the usual operator norm. For nonlinear operators $F : \mathcal{H}_1 \to \mathcal{H}_2$ we use brackets $F(v)$ whenever we write $F$ acting on $v \in \mathcal{H}_1$. Lastly, by $\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)$ we mean the space of linear and bounded operators mapping from $\mathcal{H}_1$ to $\mathcal{H}_2$, where we use the convention $\mathcal{L}(\mathcal{H}) = \mathcal{L}(\mathcal{H}, \mathcal{H})$.

II. PRELIMINARIES AND STATEMENT OF OBSERVER DESIGN PROBLEM

Let $A : D(A) \to \mathcal{H}$ be a linear operator that generates a $C_0$-semigroup $T(t)$ on $\mathcal{H}$ and $F : \mathcal{H} \times [0, t_f] \to \mathcal{H}$ be strongly continuous in time and nonlinear on $\mathcal{H}$ satisfying $F(0, t) = 0$ for every $t$.

We consider the semilinear evolution system

$$
\begin{align*}
\frac{\partial z(t)}{\partial t} &= Az(t) + F(z(t), t) + Bu(t) + G\omega(t) \\
z(0) &= z_0 \in \mathcal{H}
\end{align*}
$$

where $u(t) \in \mathcal{L}([0, t_f], \mathcal{H}_1)$ is the control input, $\omega(t) \in \mathcal{L}([0, t_f], \mathcal{H}_2)$ is the input disturbance and $B \in \mathcal{L}(\mathcal{H}_1, \mathcal{H}), G \in \mathcal{L}(\mathcal{H}_2, \mathcal{H})$. We refer to $z(t)$ as the state of the system (1).

We impose the following regularity on $F$, which is henceforth assumed to hold throughout this article.

Assumption 2.1: The operator $F$ admits a Fréchet-derivative $DF(\cdot, \cdot)$ that is globally bounded (in operator norm) as well as locally Lipschitz, uniformly in time. More precisely, there exists a constant $\delta_{DF} > 0$ such that $|DF(x, t)| \leq \delta_{DF}$ for all $(x, t) \in \mathcal{H} \times [0, t_f]$, and for every $\delta > 0$ there exists a Lipschitz
constant $\nu_{DF} > 0$ such that for all $\|x - y\| < \delta$ and all $t \in [0, t_f]$
\[ \|DF(x, t) - DF(y, t)\| \leq \nu_{DF}\|x - y\|. \]

Although the following is a well-known consequence, it is provided as proposition for the sake of completeness.

**Proposition 2.2:** The operator $F$ is globally Lipschitz, uniformly in time, meaning that there exists $M > 0$ such that for all $x, y \in \mathcal{H}$ and $t \in [0, t_f]$
\[ \|F(x, t) - F(y, t)\| \leq M\|x - y\|. \]

**Proof:** The result follows from the mean value theorem [44, Th. 5.1.12] for $M = \nu_{DF}$. □

For convenience, the disturbance $\omega$ and control $u$ may be lumped as a single input
\[ B_d = [B, G], \quad u(t) = [u^T(t), \omega^T(t)]. \]

The state-equation for $z$ in system (1) can then be written in the general form
\[ \frac{\partial z(t)}{\partial t} = Az(t) + F(z(t), t) + B_d u_d(t), \quad z(0) = z_0 \in \mathcal{H}. \] (2)

It is useful to establish in what sense the systems considered in this article admit solutions. Recall that $T(t)$ is the $C_0$-semigroup generated by $A$.

**Definition 2.3:** We say $z(t) \in C([0, t_f], \mathcal{H})$ is a mild solution of (2) if for $t \in [0, t_f]$ it satisfies the integral equation
\[ z(t) = T(t)z_0 + \int_0^t T(t - s)\left(F(z(s), s) + B_d u_d(s)\right)ds. \] (3)

It is worth noting that if $z(t)$ is a classical, that is, continuously differentiable solution to (2), then it clearly satisfies (3). However, to obtain a classical solution, one has to at least impose Lipschitz continuity in time on $F$. Such conditions are often not met in applications. The systems considered in this article are of a more general form, and the following result, ensures the existence of their mild solutions. For the proof of the following result we refer to [45, Th. 6.1.12].

**Theorem 2.4:** Consider a system of the form (2), where $A$ generates the $C_0$-semigroup $T(t)$ on $\mathcal{H}$, the nonlinearity $F(x, t)$ satisfies Assumption 2.1 and $u(t) \in C([0, t_f])$. Then, (2) has a unique mild solution $z(t) \in C([0, t_f], \mathcal{H})$ given by formula (3).

Let the system measurement be
\[ y(t) = Cz(t) + \eta(t) \]
where $\eta(t) \in C([0, t_f], \mathbb{R}^p)$, $p \geq 1$, is the output disturbance, and $C \in \mathcal{L}(\mathcal{H}, \mathbb{R}^p)$.

Our objective is to design an observer for the system (2). Most generally, an observer is a dynamical system with state $\hat{z}(t)$ such that, in the absence of disturbances
\[ \lim_{t \to \infty} \|z(t) - \hat{z}(t)\| = 0. \]

In this article, as is common, the observer dynamics contain a copy of the system’s dynamics and a feedback term that corrects for the error between the predicted observation, $C\dot{z}$, and the actual observation, $y$. The general form of the observer is
\[ \frac{\partial \hat{z}(t)}{\partial t} = A\hat{z}(t) + F(\hat{z}(t), t) + B u(t) + L(t)[y(t) - C \hat{z}(t)] \]
\[ \hat{z}(0) = \hat{z}_0 \in \mathcal{H} \] (4)

where $L(t)$, referred to as observer gain, needs to be selected so that in the absence of disturbances $\omega(t)$ and $\eta(t)$, $\hat{z}(t) \to z(t)$. The following proposition follows from Theorem 2.4 and ensures the existence of a mild solution $\hat{z}$ to (4) if $K$ is strongly continuous.

**Proposition 2.5:** Let the assumptions of Theorem 2.4 hold and let again $z(t)$ be the mild solution to (2). Let moreover $L \in C([0, t_f], \mathcal{L}(\mathbb{R}^p, \mathcal{H}))$. Then, there exists a unique mild solution $\hat{z}(t)$ to (4) given by
\[ \hat{z}(t) = T(t)\hat{z}_0 + \int_0^t T(t - s)[F(\hat{z}(s), s) + Bu(s) + F(\hat{z}(s), s) + B_d u_d(s)ds] ds. \]

**Proof:** It suffices to note that since $\eta(t) \in C([0, t_f], \mathbb{R}^p)$, also $y(t) = Cz(t) + \eta(t) \in C([0, t_f], \mathbb{R}^p)$ and hence the nonlinear operator $\hat{F}(x, t) = F(x, t) + L(t)[y(t) - Cz(t)]$ satisfies Assumption 2.1. □

The following generalization of semigroups is useful for time-varying systems.

**Definition 2.6:** Let $\Delta(t_f) := \{(t, s) : 0 \leq s \leq t \leq t_f\}$. A mapping $U(t, s) : \Delta(t_f) \to \mathcal{L}(\mathcal{H})$ is an evolution operator, if the following holds.

1) $U(t, t) = I$, $t \in \Delta(t_f)$, where $I$ denotes the identity operator.

2) $U(t, r)U(r, s) = U(t, s), 0 \leq s \leq r \leq t \leq t_f$.

3) $U(t, \cdot)$ is strongly continuous on $[s, t_f]$ and $U(\cdot, s)$ is strongly continuous on $[0, t]$.

For the proof of the following property, as well as more details on evolution operators, see for example, [46, Sec. 5.3].

**Theorem 2.7:** Let $A$ be the generator of a $C_0$-semigroup $T(t)$ on $\mathcal{H}$ and let $D(t) \in C([0, t_f], \mathcal{L}(\mathcal{H}))$. Then, there exists a unique evolution operator $U(t, s)$ satisfying
\[ U(t, s)x = T(t - s)x - \int_s^t \int_0^r T(t - r)D(r)U(r, s)xdr \] (6)
for all $x \in \mathcal{H}$. We call $U(t, s)$ the evolution operator generated by $A + D(t)$.

For $f(t) \in C([s, t_f], \mathcal{H})$, the mild solution to
\[ \frac{\partial v(t)}{\partial t} = (A + D(t))v(t) + f(t), \quad v(s) = v_0 \in \mathcal{H} \]
is
\[ v(t) = U(t, s)v_0 + \int_s^t U(t, r)f(r)dr. \]

In the next section, a method of calculating the observer gain $L(t)$ based on the EKF is described.

III. DEFINITION AND WELL-POSEDNESS OF EKF

The problem of observer design for linear systems has been well studied. The most widely known and used approach is the KF. Consider a time-varying linear system

\[ \frac{\partial z(t)}{\partial t} = \tilde{A}(t)z(t) + Bu(t) + \omega(t) \]
\[ y(t) = Cz(t) + \eta(t) \] (7)
where $\hat{A}(t)$ generates an evolution operator $U_{\hat{A}}(t,s)$ and $w(t)$ and $\eta(t)$ are process and output disturbance, respectively.

**Assumption 3.1:** Let linear operators $P_0 \in \mathcal{L}(\mathcal{H})$, $W(t) \in \mathcal{C}([0,t_f], \mathcal{L}(\mathcal{H}))$ and $R(t) \in \mathcal{C}([0,t_f], \mathcal{L}(\mathbb{R}^p))$ be self-adjoint. The operator $P_0$ is positive definite; that is, for every nonzero $v \in \mathcal{H}$, $(v,P_0v) > 0$. Moreover, for all $t \in [0,t_f]$, $W(t)$ is nonnegative definite; for every nonzero $v \in \mathcal{H}$, $(v,W(t)v) \geq 0$.

Finally, for all $t \in [0,t_f]$ the operator $R(t)$ is uniformly coercive; meaning that there exists a $\delta_0 > 0$ such that for every $w \in \mathbb{R}^p$ and $t \in [0,t_f]$, $(w,R(t)w) \geq \delta_0 \|w\|^2$.

Note that since $R(t)$ is self-adjoint, coercive and bounded for each $t$, it follows that for all $t \in [0,t_f]$, it has a self-adjoint bounded inverse $R^{-1}(t) \in \mathcal{L}(\mathcal{H})$.

Linear integral Riccati equations are defined in the following theorem. For the proof we refer to [41, Th. 3.1 and 3.3] or [47, Th. 2.1].

**Theorem 3.2:** Let $U(t,s)$ be an evolution operator on $\Delta(t_f)$. Under Assumption 3.1, the following two integral Riccati equations:

$$P(t)w = U_{P}(t,0)P_{0}U_{P}^{*}(t,0)w + \int_{0}^{t} U_{P}(t,s)W(s)U_{P}^{*}(t,s)wds$$

(8)

$$P(t)w = U_{P}(t,0)P_{0}U_{P}^{*}(t,0)w + \int_{0}^{t} U_{P}(t,s)(W(s) + P(s)C^{*}R^{-1}(s)CP(s))U_{P}^{*}(t,s)wds$$

(9)

for $w \in \mathcal{H}$, where the perturbed evolution operator $U_{P}(t,s)w = U(t,s)w - \int_{s}^{t} U(t,r)P(r)C^{*}R^{-1}(r)CU_{P}(r,s)wdr$ are equivalent and admit a unique, positive definite, and self-adjoint solution $P(t) \in \mathcal{C}([0,t_f], \mathcal{L}(\mathcal{H}))$.

Letting $U$ be the evolution operator generated by $\hat{A}$, observer dynamics for (7) are

$$\frac{\partial \hat{z}(t)}{\partial t} = \hat{A}(t)\hat{z}(t) + Bu(t) + L(t)(y(t) - C\hat{z}(t)).$$

In the case that $w(t)$ and $\eta(t)$ are process and sensor noises with covariances $W(t)$ and $R(t)$, respectively, and the covariance of the initial condition $\hat{z}(0)$ is $P_0$, then the observer gain $L(t)$ and corresponding estimate $\hat{z}(t)$ are optimal in a sense that $\hat{z}(t)$ minimizes the error covariance [47]. The observer in this case is the KF. For details, see the book [48] and for recent work on time-varying systems, [5].

The linearization of (2) will be used to define a integral Riccati equation similar to (8) and (9). Their solution $P$ define the observer gain. In the case of finite-dimensional systems, this approach is known as an EKF and this terminology will be used here.

First, the linearization of the system is defined. For this purpose, for $\hat{z}(t) \in C([0,t_f]; \mathcal{H})$, at time $t$ the Fréchet-derivative of $F(.t)$, denoted by $DF(.t) : \mathcal{H} \rightarrow \mathcal{L}(\mathcal{H})$, is

$$DF(\hat{z}(t),t) = \frac{\partial F(\hat{z}(t),t)}{\partial \hat{z}(t)} \bigg|_{\hat{z}(t)=\hat{z}(t)}.$$  

(11)

Linearizing the system (2) around $\hat{z}(t)$ yields

$$\frac{\partial \hat{z}(t)}{\partial t} = Az(t) + F(\hat{z}(t),t) + DF(\hat{z}(t),t)[\hat{z}(t) - \hat{z}(t)] + Bdu_d(t).$$

(12)

To obtain the EKF equations, the solution to a Riccati equation for the linear system (12) is needed. This Riccati equation will contain the Fréchet-derivative $DF(\hat{z}(t),t)$, a possibly nonlinear function of the observer state $\hat{z}(t)$. It is for that reason, that the Riccati equation still depends on the observer state $\hat{z}$, and recent results on the Riccati equation do not provide well-posedness for the coupled system we study here.

For a bounded linear operator $P(t) \in \mathcal{C}([0,t_f]; \mathcal{L}(\mathcal{H}))$ and $\hat{z}_0 \in \mathcal{H}$, define $\hat{z}_P(t)$ as

$$\hat{z}_P(t) = T(t)\hat{z}_0 + \int_{0}^{t} T(t-s)(F(\hat{z}_P(s),s) + Bu(s))ds + \int_{0}^{t} T(t-s)P(s)C^{*}R^{-1}(s)[y(s) - C\hat{z}_P(s)]ds.$$  

(13)

If this equation has a unique solution for $\hat{z}_P$, it defines a mapping $\hat{z}_P(t) = G_1(P(t))$. For $\hat{z}_P(\cdot) \in C([0,t_f]; \mathcal{H})$, $w \in \mathcal{H}$, an evolution operator is defined as

$$U(t,s)w = T(t-s)w + \int_{s}^{t} T(t-r)DF(\hat{z}_P(r),r)U(r,s)wdr.$$  

(14)

This evolution operator is generated by the time-varying operator $A + DF(\hat{z}_P(t),t)$. For each $\hat{z}_P(\cdot)$ the associated time-varying system, $w \in \mathcal{H}$ and the operators $P_0, R, W$, imply an integral Riccati equation

$$U_{P}(t,s)w = U(t,s)w - \int_{s}^{t} U(t,r)P(r)C^{*}R^{-1}(r)CU_{P}(r,s)wdr$$

(15)

$$P(t)w = U_{P}(t,0)P_{0}U_{P}^{*}(t,0)w + \int_{0}^{t} U_{P}(t,s)W(s)U_{P}^{*}(t,s)wds.$$  

(16)

This defines a second mapping $P(t) = G_2(\hat{z}_P(t))$. It will be shown that the mappings $G_1, G_2$ are well-defined. Then, it is shown that the composite mapping $G(\cdot) = G_2(G_1(\cdot))$ has a unique fixed point, $P(t) = G(P(t))$. This will show that the observer dynamics coupled with extended Riccati equations are well-posed in a sense that (13), (14), (15), (16) have a unique solution $\hat{z}_P(t) \in C([0,t_f]; \mathcal{H})$ and $P(t) \in C([0,t_f]; \mathcal{L}(\mathcal{H}))$. The operator

$$L(t) = P(t)C^{*}R^{-1}(t)$$  

(17)
will then define an observer (4). The initial proof of well-posedness can also be found in [42, Ch. 7].

**Proposition 3.3.** The mapping $G_1$ defined by (13) or $\dot{z}_P(t) = G_1(P(t))$ is well-defined and $G_1 : C([0, t_f], L(\mathcal{H})) \rightarrow C([0, t_f], \mathcal{H})$.

**Proof:** Due to $P(t) \in C([0, t_f], L(\mathcal{H}))$ and Assumption 3.1, $L(t) \in C([0, t_f], L(\mathbb{R}^p, \mathcal{H}))$. Hence, by Proposition 2.5 $\dot{z}_P(t) \in C([0, t_f], \mathcal{H})$.

**Proposition 3.4.** The mapping defined by (14)–(16) or $G_2 : C([0, t_f], \mathcal{H}) \rightarrow C([0, t_f], L(\mathcal{H}))$

is well-defined.

**Proof:** It suffices to note that for $x(t) \in C([0, t_f], \mathcal{H})$, $D(x(t), t) \in C([0, t_f], L(\mathcal{H}))$ and therefore, by Theorem 2.7 the operator $A + D(x(t), t)$ generates an evolution operator $U_x(t, s)$

$$U_x(t, s)w = T(t - s)w + \int_s^t T(t - r)DF(x(r), r)U_x(r, s)wdr$$

(18)

for $w \in \mathcal{H}$. By Theorem 3.2, there is a unique, positive, self-adjoint $P_x(t) \in C([0, t_f], L(\mathcal{H}))$ satisfying for all $w \in \mathcal{H}$

$$U_P(t, 0)w = U_x(t, 0)w$$

(19)

$$-\int_0^t U_x(t, r)P_x(r)C^*R^{-1}(r)CU_P(r, 0)wdr$$

$$P_x(t)w = U_P(t, 0)P_0U_x^*(t, 0)w$$

$$+ \int_0^t U_P(t, s)W(s)U_x^*(t, s)wds.$$ (20)

Therefore, the mapping $G_2(\cdot)$ is well-defined from its domain to the range.

The following result extends EKF-based observer design to the class of semilinear infinite-dimensional systems considered in this article.

**Theorem 3.5:** Let Assumptions 2.1 and 3.1 hold. For any $u(t) \in C([0, t_f], \mathcal{H})$, $P(t) \in C([0, t_f], \mathbb{R}^p)$ and $\delta_0 \in \mathcal{H}$ there exist $\delta_P(t) \in C([0, t_f], \mathcal{H})$ and $P(t) \in C([0, t_f], L(\mathcal{H}))$ such that $\dot{z}_P(t)$ solves (13) and $P(t)$ satisfies the Riccati (16), coupled to (14) and (15).

**Proof:** For $x = \dot{z}_P(t)$, define the mapping $G = G_2 \circ G_1$ by

$$G : C([0, t_f], L(\mathcal{H})) \rightarrow C([0, t_f], L(\mathcal{H}))$$

(21)

$$P(t) \mapsto P(t, \dot{z}_P(t)).$$

It will now be shown that $G$ has a unique fixed point in $C([0, t_f]; L(\mathcal{H}))$. This means that there is a unique pair $(\dot{z}_P(t), P(t))$ with $\dot{z}_P(t)$ satisfying the semilinear system (13) and $P(t)$ the Riccati (15) and (16) or (19) and (20) for $x(t) = \dot{z}_P(t)$. This will imply that the semilinear system (4) coupled with the Riccati (14)–(16) has a unique solution.

The proof is divided into the following three steps.

1) Define the closed ball

$$\mathbb{P}_{t_f}(\delta_p) = \{P(t) \in C([0, t_f]; L(\mathcal{H})), \|P(t)\| \leq \delta_p\}.$$ (22)

It will be shown that for large enough $\delta_p$, the mapping $G$ maps the ball $\mathbb{P}_{t_f}(\delta_p)$ to itself, i.e., $G : \mathbb{P}_{t_f}(\delta_p) \rightarrow \mathbb{P}_{t_f}(\delta_p)$.

2) Show $G^n$ is contractive on $\mathbb{P}_{t_f}(\delta_p)$ for large enough $n$.

3) A fixed point argument concludes the proof [49, Lemma 5.4.3].

**Step 1:** In order to prove the first part of the proof, let $\delta_{DF}$ be the bound for $Df(\cdot, \cdot)$ given by Assumption 2.1. Choose $P(t) \in C([0, t_f], L(\mathcal{H}))$. Using the Grönwall inequality, we can conclude from (18) that for all $t \in [0, t_f]$:

$$\max_{0 \leq s \leq t} \|U_{z_P}(t, s)\| \leq \delta_{T,\alpha} + \delta_{T,\alpha}\delta_{DF} \int_0^t \max_{0 \leq s \leq r} \|U_{z_P}(r, s)\|dr$$

$$\leq \delta_{T,\alpha}\exp(\delta_{T,\alpha}\delta_{DF}t_f) =: \delta_{z_P}$$ (21)

where $\delta_{T,\alpha} = \max_{t \in [0, t_f]} \|T_\alpha(t)\|$. Now, let $P(t) \in \mathbb{P}_{t_f}(\delta_p)$.

Define $\delta_W = \max_{t \in [0, t_f]} \|W(t)\|$. By [41, Lemma 2.2],

$$\max_{t \in [0, t_f]} \|G(P(t))\| \leq (\|P_0\| + t_f\delta_W\delta_{T,\alpha}) \delta_{z_P} \leq \delta_p$$

where the last inequality is true if $\delta_p$ is sufficiently large. For such $\delta_p$, we can conclude the first part of the proof, or

$$G : \mathbb{P}_{t_f}(\delta_p) \rightarrow \mathbb{P}_{t_f}(\delta_p).$$

**Step 2:** Now, we prove that $G^n$ is a contraction on $\mathbb{P}_{t_f}(\delta_p)$ for large enough $n \in \mathbb{N}$. For $i = 1, 2$, let $P_i(t) \in \mathbb{P}_{t_f}(\delta_p)$ and define in the following:

1) $\dot{z}_{P,i}(t) = G_1(P_i(t))$ satisfying (13) with $P(t) = P_i(t)$.

2) $U_{\dot{z}_{P,i},i}(t, s)$ being the evolution operator given by (18) with $x(t) = \dot{z}_{P,i}(t)$.

3) $U_{\dot{z}_{P,i},i}(t, s)$ being the perturbation of $U_{\dot{z}_{P,i}}(t, s)$ by $-G(P_i(t))C^*R^{-1}(t)$C given by (19).

Moreover define

$$\Delta \dot{z}_P(t) = \dot{z}_{P,1}(t) - \dot{z}_{P,2}(t)$$

$$\Delta U_{\dot{z}_P}(t, s) = U_{\dot{z}_{P,1},1}(t, s) - U_{\dot{z}_{P,2},2}(t, s)$$

$$\Delta P(t) = P_{1}(t) - P_2(t)$$

$$\Delta U_P(t, s) = U_{P,1}(t, s) - U_{P,2}(t, s)$$

$$\Delta G(t) = G(P_1(t)) - G(P_2(t)).$$

**Step 2.1:** In this step, it is shown that the operators $\Delta U_P(t, s)$ and $\Delta U_{\dot{z}_P}(t, s)$ are bounded and the bounds are defined. For $\Delta \dot{z}_P(t)$ satisfying (13), by Assumption 2.1 and the Grönwall inequality, it can be shown that there exists $\Delta \dot{z}_P > 0$ such that for all $t \in [0, t_f]$

$$\|\Delta \dot{z}_P(t)\| \leq \delta_{z_P}, \text{ for all } P(t) \in \mathbb{P}_{t_f}(\delta_p).$$

(22)

For the difference $\Delta \dot{z}_P$ we compute

$$\Delta \dot{z}_P(t) = \int_0^t T(t - s)\left(F(\dot{z}_{P,1}(s), s) - F(\dot{z}_{P,2}(s), s)\right)ds$$

$$+ \int_0^t T(t - s)\Delta P(t)C^*R^{-1}(s)g(s)ds$$

$$- \int_0^t T(t - s)\left(\Delta P(t)C^*R^{-1}(s)C\dot{z}_{P,1}(s)

+ P_2(s)C^*R^{-1}(s)C\Delta \dot{z}_P(s)\right)ds.$$ (23)

From (23), Assumptions 2.1 and 3.1, inequality (22), and boundedness of the operators $T(t - s), C$, and $R^{-1}(s)$, it is concluded that there are constants $\delta_1, \delta_2 > 0$ and $c_1 > 0$ such that for all
where the last inequality was obtained by applying the Grönwall inequality.

Similarly, from (18), it is derived that for all \( w \in \mathcal{H} \)
\[
\Delta U_{z_p}(t, s) = \frac{1}{s} \int_{t}^{s} \frac{1}{r} \int_{s}^{r} \Delta U_{z_p}(t, r) \Delta U_{z_p}(r, s) \, w \, dr \, ds
\]
\[+ \int_{s}^{t} \frac{1}{r} \int_{s}^{r} \Delta U_{z_p}(t, r) \Delta U_{z_p}(r, s) \, w \, dr \, ds.
\]

Given Assumption 2.1, the boundedness defined by (22), and the boundedness of the operators \( T(t - s) \) and \( U_{z_p,1}(r, s) \) defined by (21), (25) allow us to compute
\[
\max_{0 \leq s \leq t} \| \Delta U_{z_p}(t, s) \| \leq \delta U_{z_p} \int_{0}^{t} \| \Delta U_{z_p}(s) \| \, ds
\]
\[+ \frac{1}{c} \max_{0 \leq r \leq s} \| \Delta U_{z_p}(r, s) \| \, dr.
\]

which by Grönwall inequality yields for all \( t \in [0, t_f] \)
\[
\| \Delta U_{z_p}(t, s) \| \leq c_2 \int_{0}^{t} \| \Delta U_{z_p}(s) \| \, ds
\]
for some \( c_2 > 0 \). From (24) and (26), it is concluded that for all \( t \in [0, t_f] \)
\[
\| \Delta U_{z_p}(t, s) \| \leq c_2 c_1 \int_{0}^{t} \| \Delta P \| \, ds.
\]
Note that by the boundedness of \( U_{z_p,1}(t, s) \), \( i = 1, 2 \) given by (21), as well as Grönwall’s lemma, it can be shown that the perturbed evolution operators \( U_{p,i}(t, s) \), \( i = 1, 2 \) are, for all \( t \in [0, t_f] \), also bounded by
\[
\max_{0 \leq s \leq t} \| U_{p,i}(t, s) \| \leq \delta U_{p}
\]
for some \( \delta U_{p} > 0 \).

Let \( Q(t) = C^* R^{-1}(t) C \). Similarly, from (19), it can be derived that for all \( w \in \mathcal{H} \)
\[
\Delta U_{P}(t, s)w = \frac{1}{s} \int_{t}^{s} \frac{1}{r} \int_{s}^{r} \Delta U_{P}(t, r) Q(r) U_{P,1}(r, s) \, w \, dr \, ds
\]
\[+ \int_{s}^{t} \frac{1}{r} \int_{s}^{r} \Delta U_{P}(t, r) Q(r) U_{P,1}(r, s) \, w \, dr \, ds.
\]

Using the boundedness defined by (21) and (28) as well as the boundedness of the operator \( Q(t) \), by Grönwall’s inequality, we can show that (29) leads to
\[
\max_{0 \leq s \leq t} \| \Delta U_{P}(t, s) \| \leq c_3 \max_{0 \leq s \leq t} \| \Delta U_{z_p}(t, s) \|
\]+ \[+ c_4 \int_{0}^{t} \| \Delta G(r) \| \, dr.
\]
for some constants \( c_3, c_4 > 0 \); substituting (27) into this inequality yields for all \( t \in [0, t_f] \)
\[
\| \Delta U_{P}(t, s) \| \leq c_4 \int_{0}^{t} \max_{0 \leq r \leq s} \| \Delta P(r) \| \, ds + c_4 \int_{0}^{t} \| \Delta G(r) \| \, dr
\]
(30)

where \( c_4 = c_3 c_2 c_1 \).

Step 2.2: This step, we use the bounds obtained in the above to find a bound for \( G(P_1(t)) - G(P_2(t)) \). Note that the operators \( G(P_i(t)) \) for \( i = 1, 2 \) satisfy
\[
G(P_1(t)) = U_{P,1}(t, 0) P_1 U_{z_p,1}^*(t, 0) w
\]+ \[+ \int_{0}^{t} U_{P,1}(t, s) W(s) U_{z_p,1}^*(t, s) \, w \, ds
\]
(31)

for all \( w \in \mathcal{H} \). From (31), it can be concluded that the difference \( G(P_1(t)) - G(P_2(t)) \) satisfies, for all \( w \in \mathcal{H} \)
\[
(G(P_1(t)) - G(P_2(t))) \, w = \frac{1}{s} \int_{t}^{s} \frac{1}{r} \int_{s}^{r} \Delta U_{P}(t, r) Q(r) U_{P,1}(r, s) \, w \, dr \, ds
\]
\[+ \int_{s}^{t} \frac{1}{r} \int_{s}^{r} \Delta U_{P}(t, r) Q(r) U_{P,1}(r, s) \, w \, dr \, ds
\]
\[+ U_{P,1}(t, s) W(s) U_{z_p,1}^*(t, s) \, w \, ds
\]
which given (21) and (28) as well as the boundedness of the operators \( P_i \) and \( W(t) \) leads to
\[
\| G(P_1(t)) - G(P_2(t)) \| \leq c_5 \max_{0 \leq r \leq t} \| \Delta U_{z_p}(t, r) \|
\]
\[+ c_4 \int_{0}^{t} \| \Delta G(r) \| \, dr.
\]
(32)

for all \( t \in [0, t_f] \) and for some \( c_5 > 0 \).

Substituting (27) and (30) into (32) results in
\[
\| G(P_1(t)) - G(P_2(t)) \| \leq c_5 \int_{0}^{t} \max_{0 \leq r \leq s} \| \Delta P(r) \| \, ds
\]
\[+ c_4 \int_{0}^{t} \| \Delta G(s) \| \, ds
\]
where \( c_5 = (c_5 c_2 c_1)(c_5 + 1) \), for all \( t \in [0, t_f] \), and employing Grönwall’s inequality yields for all \( t \in [0, t_f] \),
\[
\| G(P_1(t)) - G(P_2(t)) \| \leq c \int_{0}^{t} \max_{0 \leq r \leq s} \| P_1(r) - P_2(r) \| \, ds
\]
(33)

for some \( c > 0 \).

Now, we show that \( G^n(\cdot) \) is a contraction mapping for large enough \( n > 0 \). For this purpose, we use an induction argument to show that
\[
\| G^n(P_1(t)) - G^n(P_2(t)) \| \leq (c t)^n \max_{0 \leq r \leq t} \| P_1(r) - P_2(r) \|.
\]
(34)
First, for $n = 1$ the argument holds by (33). Now let $n \geq 2$ and assume (34) holds for $k \leq n - 1$. By (33)
\[
\|G^n(P_1(t)) - G^n(P_2(t))\| \\
\leq c \int_0^t \max_{\theta \in \mathbb{T}_s} \|G^{n-1}(P_1(\tau)) - G^{n-1}(P_2(\tau))\| d\tau \\
\leq \frac{c^n}{(n-1)!} \int_0^t \max_{\theta \in \mathbb{T}_s} \|P_1(\tau) - P_2(\tau)\| d\tau \\
\leq \frac{(c)t^n}{n!} \max_{\theta \in \mathbb{T}_s} \|P_1(\tau) - P_2(\tau)\|
\]
which proves (34). By taking the maximum on $[0, t_f]$ it follows that
\[
\max_{[0,t_f]} \|G^n(P_1(t)) - G^n(P_2(t))\| \leq \frac{(c)t^n}{n!} \max_{[0,t_f]} \|P_1(t) - P_2(t)\|.
\]
Therefore, for $n$ large enough $(c)t^n < 1$ and $G^n(\cdot)$ is a contraction on $\mathbb{P}_{t_f}((\delta_p))$. Thus, by the contraction mapping theorem, see for example, [49, Lemma 5.4-3], there is a unique fixed point on $\mathbb{P}_{t_f}((\delta_p))$, completing the proof.

This implies the existence of a unique mild solution of the observer dynamics (4) with observer gain $L(t)$ defined by (17) where the linear operator $P(t)$ is the solution of the Riccati (14)–(16). In order to adjust the rate at which the estimation error converges to zero, in the Riccati coupled (14)–(16), the operator $A$ is replaced by $A + \alpha I$ where $\alpha > 0$ and $I$ is the identity operator. In other words, let $T_\alpha(t)$ be the $C_0$-semigroup generated by $A + \alpha I$ and for $z_\nu(t) \in C([0, t_f]; \mathcal{H})$ and $w \in \mathcal{H}$ defines a evolution operator
\[
U_\alpha(t, s)w = T_\alpha(t - s)w \\
+ \int_s^t T_\alpha(t - r)DF(z_\nu(r), r)U_\alpha(r, s)wdr.
\]

The integral Riccati equation takes the form
\[
U_{P(\alpha)}(t, s)w = U_\alpha(t, s)w \\
- \int_s^t U(t, r)P(r)C^*R^{-1}(r)CU_{P(\alpha)}(r, s)wdr \\
P(t)w = U_{P(\alpha)}(t, 0)P_0U_{P(\alpha)}(t, 0)w \\
+ \int_s^t U_{P(\alpha)}(t, s)(W(s) \\
+ P(s)C^*R^{-1}(s)CP(s))U_{P(\alpha)}(t, s)wds.
\]

Note that (35)–(37) satisfy the conditions of Theorem 3.5 and, thus, are well-posed on every bounded time interval $[0, t_f]$.

IV. STABILITY OF THE LINEARIZED SYSTEM

Our next main objective is to show uniform convergence of the estimator state to the system true state under some conditions. For this purpose, it is essential to first show the exponential stability of the evolution operator generated by the time varying linear operator $A + DF(z_\nu(t), t) - L(t)C$ where $L(t) = P(t)C^*R^{-1}(t)$ and $P(t)$ solves (35)–(37). This is the purpose of this section. In order to obtain such stability, existing results for finite-time linear quadratic control needed to be extended. In order to maintain the flow of this article, these results are in an appendix at the end of this article. Duality between the linear time-varying operator $A + DF(z_\nu(t), t) - L(t)C$ and the adjoint operator defined by a dual control problem is used to prove the stability of the linearized system in this section.

The following definition is standard.

Definition 4.1: An evolution operator $Y(t, s)$ is exponentially stable if there exists $M \geq 0$, $\alpha > 0$ such that for all $t \geq s$
\[
\|Y(t, s)\| \leq Me^{-\alpha(t-s)}.
\]

Let $A_c(t) : \mathcal{D}(A_c(t))\mathcal{H} \to \mathcal{H}$ be a linear operator that generates an evolution operator $Y(t, s)$ and consider bounded linear operators $B_c(t) \in \mathcal{C}([0, \infty); \mathcal{L}(\mathcal{X}_B, \mathcal{H}))$ and $C_c(t) \in \mathcal{C}([0, \infty); \mathcal{L}(\mathcal{H}, \mathcal{X}_C))$ where $\mathcal{X}_B$ and $\mathcal{X}_C$ are Hilbert spaces.

Definition 4.2:

1) For a linear bounded and continuous in time operator $L(t)$ define the evolution operator $Y_L(t, s)$ satisfying
\[
Y_L(t, s)w = Y(t, s)w \\
+ \int_s^t Y(t, r)L(r)C_n(r)Y_L(r, s)wdr
\]
for $w \in \mathcal{H}$. The pair $(A_c(t), C_c(t))$ is uniformly detectable if there exists a linear uniformly bounded and continuous in time operator $L(t)$ such that the evolution operator $Y_L(t, s)$ is exponentially stable.

2) For any linear uniformly bounded and continuous in time operator $K(t)$ define the evolution operator $Y_K(t, s)$ satisfying
\[
Y_K(t, s)w = Y(t, s)w \\
+ \int_s^t Y(t, r)B_n(r)K(r)Y_K(r, s)wdr
\]
for $w \in \mathcal{H}$. The pair $(A_c(t), B_c(t))$ is uniformly stabilizable if there exists a linear uniformly bounded and continuous in time operator $K(t)$ such that the evolution operator $Y_K(t, s)$ is exponentially stable.

The stability of the linearized error dynamics follows from the following theorem.

Theorem 4.3: Defining $A_d(t) = A + \alpha I + DF(z_\nu(t), t)$, assume that $(A_d(t), C)$, is uniformly detectable and $(A_d(t), W^{1/2}(t))$ is uniformly stabilizable. Defining $L(t) = P(t)C^*R^{-1}(t)$ where $P(t)$ satisfies the Riccati (37), the evolution operator $Y_d(t, s)$ generated by $A_d(t) - L(t)C$ is uniformly bounded; that is there exists $\delta_\nu > 0$ can be chosen so that for all $t \in [0, t_f]$
\[
\|Y_d(t, s)\| \leq \delta_\nu.
\]

Proof: A standard duality argument will be used to show the stability of the evolution operator generated by $A_d - L(t)C$.

Let $A_d(t) = A_d^0(t_f - t)$, $B_c(t) = C^*$, $W_c(t) = W(t_f - t)$, and $R_c(t) = R(t_f - t)$. $Y(t, s) = U_{P(\alpha)}(t_f - s, t_f - t)$, where $U_{P(\alpha)}(t, s)$ is generated by $A_d(t)$; thus, $Y(t, s)$ is a mild evolution operator. For any $z_0 \in \mathcal{H}$, define
\[
P(t_f)(0)z_0 = Y_{-B_c^*, K}^{+}(t_f, 0)P_0Y_{-B_c^*, K}^{-}(t_f, 0)z_0
\]
\[
Y_{-B_{t}}(t, s) = Y(t, s) \z_0 - \int_{s}^{t} Y(t, r)B_{c}(r)R_{c}(r)B_{c}^{\dagger}(r)P_{t}(r)Y_{-B_{t}}(r, s) \z_0 dr.
\]

Finally, the result now follows from duality; since \( U_{P,\alpha}(t, s) = Y_{-B_{t}}(t, s) \) by [41, Th. 1.2 and 1.3] and definition of \( Y_{-B_{t}}(t, s) \), (40) is preserved.

**Lemma 4.4:** For \( D_{0}(t) \in \mathbb{L}(\mathcal{H}, \mathcal{H}) \) continuous in time define \( U_{D_{0}}(t, s) \) to be the evolution operator generated by \( A + D_{0}(t) \). Furthermore, for \( \beta_{0} > 0 \), define \( U_{D_{0}}(t, s) \) to be the evolution operator generated by \( A + D_{0}(t) + \beta_{0}I \). Then
\[
U_{D_{0}}(t, s) = \exp(\beta_{0}(t - s))U_{D_{0}}(t, s).
\]

**Proof:** Note that the evolution operator \( U_{D_{0}}(t, s) \) satisfies
\[
U_{D_{0}}(t, s) = T_{\beta_{0}}(t - s)\]
\[
+ \int_{s}^{t} T_{\beta_{0}}(t - r)D_{0}(r)U_{D_{0}}(r, s) dr
\]
which \( T_{\beta_{0}}(t) \) is the semigroup generated by \( A + \beta_{0}I \). Also, from semigroup properties, it can be concluded that
\[
T_{\beta_{0}}(t - s) = \exp(\beta_{0}(t - s))T(t - s)
\]
substituting this equality back in (44) and multiplying both sides by \( \exp(-\beta_{0}(t - s)) \) result in
\[
\exp(-\beta_{0}(t - s))U_{D_{0}}(t, s) = T(t - s)
\]
\[
+ \int_{s}^{t} T(t - r)D_{0}(r)\exp(-\beta_{0}(r - s))U_{D_{0}}(r, s) dr.
\]
From definition of the evolution operator \( U_{D_{0}}(t, s) \) and (45) it is concluded that \( U_{D_{0}}(t, s) = \exp(-\beta_{0}(t - s))U_{D_{0},\beta_{0}}(t, s) \) and the proof is complete.

**V. ERROR DYNAMICS**

In this section, the convergence of the estimated state \( \hat{z}_{P}(t) \) to the true state \( z(t) \), with a filter \( L(t) \) defined by (35)–(37) as \( t_{f} \to \infty \) is shown to hold under some additional assumptions, if the initial error is sufficiently small.

Define the error \( e(t) = z(t) - \hat{z}_{P}(t) \) between the system state \( z(t) \) and the observer state \( \hat{z}_{P}(t) \) and for \( z, \hat{z}_{P} \in \mathcal{H} \)
\[
\phi(e, t) = F(z, t) - F(z - e, t) - DF(z - e, t)(e). \tag{46}
\]

From definition of the system (1) and the observer (4) with \( L(t) = P(t)C^{\dagger}R^{-1}(t) \), the differential equation governing the error dynamics is
\[
\frac{\partial e(t)}{\partial t} = A_{e}(t) - L(t)e(t) + D_{F}(z(t) - e(t), t)e(t) + \phi(e(t)) - L(t)e(t) - G\omega(t).
\]

Let \( U_{P}(t, s) \) be the evolution operator generated by \( A + D_{F}(z_{P}(t), t) - L(t)C_{e} \) where \( L(t) = P(t)C^{\dagger}R^{-1}(t) \) and \( P(t) \) solves (35)–(37). The mild solution for the error dynamics is, with initial condition \( e(0) = e_{0} \)
\[
e(t) = U_{P}(t, 0)e_{0} + \int_{0}^{t} U_{P}(t, r)\phi(e(r)) dr - \int_{0}^{t} U_{P}(t, r)(-L(r)e(r) - G\omega(r)) dr.
\]

The right-hand side of (48) defines a nonlinear mapping
\[
\Phi_{e}(t, 0, e_{0}) : L_{2}((0, t_{f}); \mathcal{H}) \to L_{2}((0, t_{f}); \mathcal{H}).
\]

Well-posedness of the differential (1) and (4) for \( z(t) \) and \( \hat{z}_{P}(t) \) implies that for every initial condition \( e(0) \in \mathcal{H} \) the above equation has a unique solution in \( C([0, t_{f}], \mathcal{H}) \) for e. Thus, for all \( e(0) \in \mathcal{H} \), the operator \( \Phi_{e} \) has a unique fixed point \( e \in L_{2}((0, t_{f}); \mathcal{H}) \).

It will be shown that under stabilizability and detectability assumptions, and in the absence of disturbances, the estimation error, defined by (47) with \( L(t) = P(t)C^{\dagger}R^{-1}(t) \) where \( P(t) \) satisfies the Riccati (37), converges to zero if the initial error is sufficiently small. A second-order smoothness assumption on the nonlinear term \( F(\cdot) \) is needed.

**Theorem 5.1:** Let the system \( (A_{d}(t), C) \), where \( A_{d}(t) = A + \alpha I + D_{F}(z_{P}(t), t) \), be uniformly detectable, and \( (A_{d}(t), W_{1/2}(t)) \) is uniformly stabilizable. Assume that there exist time \( T > 0 \) and positive numbers \( m > 1, \epsilon_{o} > 0 \) and \( \delta_{p} \) such that if \( ||e||_{\mathcal{H}} < \epsilon_{o} \), then defining \( \phi(e) \) as in (46)
\[
||\phi(e, t)||_{\mathcal{H}} \leq \delta_{p}||e||_{\mathcal{H}}, \quad 0 \leq t \leq T.
\]

From this, the convergence of the estimated state \( \hat{z}_{P}(t) \) to the true state \( z(t) \), with a filter \( L(t) \) defined by (35)–(37) as \( t_{f} \to \infty \) is shown to hold under some additional assumptions, if the initial error is sufficiently small.

Define the error \( e(t) = z(t) - \hat{z}_{P}(t) \) between the system state \( z(t) \) and the observer state \( \hat{z}_{P}(t) \) and for \( z, \hat{z}_{P} \in \mathcal{H} \)
\[
\phi(e, t) = F(z, t) - F(z - e, t) - DF(z - e, t)(e).
\]

From (48), if \( ||e_{0}||_{\mathcal{H}} \leq \epsilon \) and \( e \in S \)
\[
||\Phi_{e}(t, 0, e_{0})||_{\mathcal{H}} \leq ||U_{P}(t, 0)||_{\mathcal{H}} ||e_{0}||_{\mathcal{H}} + \int_{0}^{t} ||U_{P}(t, r)||_{\mathcal{H}} ||\phi(z(r), \hat{z}_{P}(r))||_{\mathcal{H}} dr \leq \delta_{Y} \exp(-\alpha t) ||e(0)||_{\mathcal{H}}
\]
Choose \( \epsilon > 0 \) and \( \epsilon_S > 0 \) so that
\[
\epsilon_S^{-1} < \frac{\alpha}{\delta_Y \delta_\phi}, \quad \epsilon < \frac{1}{\delta_Y} \epsilon_S.
\]
Then, for all \( 0 \leq t \leq t_f \), \( \Phi_c(t, 0, e_0) \) maps \( \mathcal{S} \) to itself. Thus, for all initial errors \( \|e(0)\| < \epsilon \), it follows that
\[
\epsilon < \frac{1}{\delta_Y} \left( \frac{\alpha}{\delta_Y \delta_\phi} \right)^{\frac{1}{m-1}}
\]
the error \( \|e(t)\| \) is uniformly bounded by some number \( \epsilon_S < \frac{\alpha}{\delta_Y \delta_\phi} \) for all \( t \).

It will now be shown that if \( \|e(0)\| < \epsilon \), the estimation error decays exponentially. Define \( \hat{e}(t) = e^{\alpha t} e(t) \). From (48) and (50)
\[
\|\hat{e}(t)\|_H \leq \delta_Y \|\hat{e}(0)\|_H + \int_0^t \delta_Y \delta_\phi \|e(r)\|^{m-1}_H \|\hat{e}(r)\|_H dr.
\]
By Gronwall’s Inequality
\[
\|\hat{e}(t)\|_H \leq \delta_Y \|\hat{e}(0)\|_H \exp(\delta_Y \delta_\phi \epsilon_S^{-m-1} t)
\]
and so
\[
\|e(t)\|_H \leq \delta_Y \|e(0)\|_H \exp \left( \frac{\epsilon_S^{-m-1} \delta_Y \delta_\phi}{\alpha} - 1 \right) t.
\]
Since \( \epsilon_S^{-m-1} < \frac{\alpha}{\delta_Y \delta_\phi} \), the error decays exponentially.

Thus, in the absence of disturbances, there is exponential convergence of the estimation error to zero, if the initial error is small enough. As expected, with smaller initial error \( \epsilon \) the exponential decay improves.

If disturbances are present, the estimation error is bounded.

**Corollary 5.2:** Consider the same assumptions as in Theorem 5.1 except that there is a disturbance \( \omega(t) \), \( \eta(t) \) satisfying for some \( \delta_\phi > 0 \) \( \|\omega(t)\|, \|\eta(t)\| \leq \delta_\phi \). There exists \( \delta_\phi > 0 \) such that if \( \|e(0)\| \leq \delta_\phi \) and the estimation error is bounded.

**Proof:** Since the undisturbed system is locally exponentially stable around the equilibrium point \( e(t) = 0 \), it is uniformly asymptotically stable in sense of [50, Def. 2.24]. Furthermore, the nonlinear part of the error dynamics defined by \( \phi(z(t), \hat{z}(t)) - K(t) \eta(t) - G \omega(t) \) is uniformly Lipschitz continuous with respect to its arguments. Therefore, by [50, Th. 2.25], there exist \( \delta_\phi > 0 \) such that for \( \|e(0)\| \leq \delta_\phi \) the disturbed system is locally input-to-state stable, and thus the estimation error \( e(t) \) is bounded.

Although it is not always straightforward to establish input uniform detectability, this condition is less restrictive than the observability condition in [20], [21], and [24]. For some \( \alpha > 0 \) consider a simple ordinary differential equation (ODE) system
\[
A(t) = \begin{bmatrix}
a_1(t) & 0 & 0 \\
0 & a_2(t) & 0 \\
0 & 0 & a_3(t) \\
\end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}
\]
with \( a_1(t) < -\alpha, a_2(t) < -\alpha \) and \( a_3(t) < m \) for some constant \( m \). Choosing \( L = \begin{bmatrix} 0 & 0 & -(m + \alpha) \end{bmatrix} \), \( \|L(t)\| \) is uniformly bounded (in fact constant). Furthermore, defining
\[
Q = \text{diag} \{-2a_1(t), -2a_2(t), 2(-a_3(t) + m + \alpha)\}
\]

A + LC generates an exponentially stable operator with decay rate exceeding \( \alpha \). Therefore, the system is uniformly detectable; however, a quick calculation reveals that this system is unobservable.

Furthermore, the uniform observability condition implies exact observability for infinite-dimensional systems. Exact observability is a very restrictive condition for infinite-dimensional systems when bounded observation is involved. For example, even with the identity as observation operator, the heat equation is not exactly observable; it is however exponentially detectable [51, Example 4.1.15].

**VI. EXAMPLE: ESTIMATION IN A MAGNETIC DRUG DELIVERY SYSTEM**

A simple magnetic drug delivery system is considered where the magnetic stimulation is utilized to control the distribution of the magnetic drug infused nanoparticles. In the first simulation, the system is simplified such that the assumptions of the theorem stated above is satisfied. Next, we consider a more general case and show that even for a less smooth nonlinearity than assumed, the observer leads to converging estimates.

The distribution \( e(r, s, t) \) of the magnetic nanoparticles is manipulated in a fluid environment via magnetic force generated by electromagnets. Details of the system can be found in [52] and also [53], [54]: only an overview is provided here. The electromagnets’ currents are represented by \( I_1(t) \) and \( I_3(t) \) in the \( r \)-direction and \( I_2(t) \) and \( I_4(t) \) in the \( s \)-direction. The currents of the Helmholtz coils are denoted by \( I_5(t) \) in the \( r \)-direction and by \( I_6(t) \) in the \( s \)-direction. The system model is [55]
\[
\dot{c}(t) = -\nabla \cdot (-D \nabla c(t) + \kappa c(t)V_f(t) + \gamma c \nabla (H^T(t)H(t)))
\]
on the domain of interest \( \Omega = [-L_0, L_0] \times [-L_0, L_0] \), where \( D \) is the diffusion coefficient, \( \kappa \) is the advection coefficient, \( \gamma \) is a coefficient defined by the magnetic properties and size of the nanoparticles, \( H \) is the magnetization vector (which is a linear function of the currents \( I_c(t) \)), \( V_f(t) \) is the flow velocity field. The equations are solved with homogeneous Neumann boundary conditions. The magnetization vector is \( H(t) = J_c(t) \) where \( I(t) = [I_1, \ldots, I_6]^T \) and \( J_c \) is a matrix defined by the configuration and magnetic characteristics of the electromagnets and Helmholtz coils (for more details see [54, Ch. 4]). Thus, \( \nabla (H^T(t)H(t)) = Q_e I_c(t) \) where \( Q_e \) is a matrix with \( i \)-th component defined by \( \nabla (J_{c,i} \cdot J_{c,j}) \) where \( J_{c,k} \) is the \( k \)-th column of \( J_c \) and

\[
I_c(t) = [I_1(t) I_1(t) I_2(t), \ldots, I_6(t)]^T.
\]

The behavior of the system (53) is simulated over a square \( \Omega \) of size \( 2 \text{ cm} \times 2 \text{ cm} \). The diffusion coefficient is \( D = 1 \times 10^{-8} \text{ m}^2/\text{s} \), \( \kappa = 2.5 \times 10^{-7} \) and \( \gamma = 6.6 \times 10^{-5} \). The particles have radius 500 nm.
is linearized around a fixed distribution $V_c(0)$. The last term on the right-hand side of (53) is linearized around a fixed distribution $c_r(r,s) = c_r \exp(-r^2 + s^2)/(6.25 \times 10^{-5})$, where $c_r > 0$ is a normalization factor defined by

$$c_r = \mathcal{A}(\Omega) / \int_{\Omega} \exp(-r^2 + s^2)/(6.25 \times 10^{-5}) dr ds$$

where $\mathcal{A}(\Omega)$ is the area of the domain $\Omega$. The governing equations now are

$$\dot{c}(t) = -\nabla \cdot (-D \nabla c(t) + \gamma c c_L c_I(t)) - \kappa c^2(t) + \omega(t).$$

We choose the external input to be

$$I(t) = [0.1 \times 9, 0.8 \sin(20t)/20, 0_1 \times 4, 16 \sin(40t)/40, 0_1 \times 30]^T.$$ 

The initial concentration is set to be uniform, $c(\cdot, \cdot, 0) \equiv 1$, and the initial condition of the estimator is set at zero.

The state space $\mathcal{H} = L^2(\Omega)$ and the state is $z(t) = c(t) \in L^2(\Omega)$. Also

$$A_c(t) = \nabla \cdot (D \nabla z(t)), \quad F(z(t)) = -\kappa z^2(t)$$

$$Bu(t) = \gamma c \nabla Q_c u(t), \quad u(t) = I(t), \quad y(t) = C z(t).$$

The operator $A$ generates a $C_0$-semigroup. The nonlinear term in (55), $F(z(t))$ is Lipchitz continuous and Fréchet differentiable but the derivative is not uniformly bounded over the entire state space. However, by [56, p. 17], uniform boundedness of the filtering gain $L(t)$, and continuity of input signals $u(t)$, the solution to both system and observer are in $H^2(\Omega)$; therefore, the concentration is continuous over $\Omega$ and $z(t) = c(t)$ is bounded over any finite interval of time. Thus, this function can be replaced by a uniformly bounded function in the analysis. Furthermore, we can show that the linear system $(A, C)$ is $\beta$-detectable for some $\beta > 0$ according to [51, Th. 5.2.7]. Also since $\kappa$ is relatively small and the concentration remains bounded, we assume that the perturbation due to the additional term $-2\kappa z_c(t)$ is small and so the linearized time varying system $(A_d(t), C)$ is uniformly detectable for some $\alpha > 0$.

Both system and observer dynamics are approximated using the finite element method with square elements and piecewise linear basis functions. The order of approximation for the system is $35 \times 35$. Four different orders of approximation were used...
for observer dynamics, \(25 \times 25, 18 \times 18, 9 \times 9\), and \(7 \times 7\). The equations are solved in MATLAB 2018. The filtering parameters are chosen to be \(\alpha = 8\), \(W(s) = I_H\), and \(R(s) = 100I_Y\) where \(I_H\) and \(I_Y\) are the identity operators on \(H\) and \(Y\), respectively. It is straightforward to prove that the system \((\mathcal{A}, W^{1/2}(t))\) with \(W = I_H\) is \(\beta\)-detectable; followed by smallness of \(\kappa\) and boundedness assumption of the concentration, we can conclude that \((\mathcal{A}_d(t), W^{1/2}(t))\) is uniformly stabilizable. The \(L^2(\Omega)\)-norm of the state estimation error and also the Euclidean norm of the error in the predicted measurement are shown in Fig. 1. The errors converge quickly to small values. As expected, the steady-state estimation error increases as the order of the observer decreases.

Simulation results in the presence of system and measurement disturbance are shown in Fig. 2. The disturbance signal \(\omega(t) \in \mathbb{R}\) is generated via MATLAB random signal generator such that it has zero mean and covariance of 0.1. The output disturbance \(\eta(t) \in \mathbb{R}^3\) is similarly created with zero mean and a covariance matrix of \(\text{diag}(5 \times 10^{-3}I_{12}, 5 \times 10^{-4})\). The state estimation error increases slightly but remains bounded.

To investigate a less smooth nonlinearity, the velocity field is defined as a function of distribution \(V_f(t) = I_{2 \times 1}c(t)\). The system (53) is reformulated as

\[
\begin{align*}
\dot{c}(t) &= -\nabla \cdot (-D\nabla c(t) + \kappa c^2(t)I_{2 \times 1} + \kappa cQ_cI_c(t)) + \omega_1(t) \\
\dot{I}_c(t) &= U_f(t) + \omega_2(t). 
\end{align*}
\]

In this new form, \(I_c(t)\) is an augmented state so that the state representation of the system follows the form given by (1). Here, \(U_f(t)\) is the derivative of current vector and is the input to the system; we chose \(U_f = [0\times 9, 0.8\cos(20t), 0\times 4, 1.6\cos(40t), 0\times 30]^T\). The state space is now \(\mathcal{H} = L^2(\Omega) \times \mathbb{R}^m\) where \(m > 0\) is the dimension of vector \(I_c\). The state is \(z(t) = (z_1(t), z_2(t), c(t), I_c(t)) \in L^2(\Omega) \times \mathbb{R}^m\). Also

\[
A_z(t) = \begin{bmatrix} \nabla \cdot (D\nabla z_1(t)) & 0 \\ 0 & 0 \end{bmatrix},
\]

\[
F(z(t)) = -\nabla \cdot (\kappa z_1^2(t)I_{2 \times 1} + \gamma z_1(t)Q_cz_2(t))
\]

\[
Bu(t) = [0, I]^T u(t), \quad u(t) = U_f(t), \quad y(t) = Cz(t).
\]

The initial conditions and other details are the same as in the previous example. The error in both the state estimation and in the measurement are shown in Fig. 3. Although the assumption of Theorem 3.5 on the system nonlinearity is not satisfied, the estimate converges to the true state in these simulations for large-order estimator, and remains small for the lower-order estimators.

In Fig. 4, the observer performance in presence of system and output disturbance is shown. The disturbance signal \(\omega_1(t) \in \mathbb{R}\) and \(\omega_2(t) \in \mathbb{R}^3\) are generated via MATLAB random signal...
subject to the dynamics
\[
\frac{\partial z_{\text{adj}}(t)}{\partial t} = A_c(t)z_{\text{adj}}(t) + B_c(t)u(t), \quad z_{\text{adj}}(0) = z_{\text{adj},0} \tag{58}
\]
with mild solution
\[
z_{\text{adj}}(t) = Y(t,0)z_{\text{adj},0} + \int_0^t Y(t,r)B_c(r)u(r)dr. \tag{59}
\]

**Theorem A.1** [41, Th. 2.1–2.3]: For any \(z_0 \in H\), the integral (41) and (42), defined by
\[
P_{t_f}(0)z_0 = Y_{-B_cK_{t_f}}(t_f,0)P_0Y_{-B_cK_{t_f}}(t_f,0)z_0 \\
+ \int_0^{t_f} Y_{-B_cK_{t_f}}(r,0)(W_c(r) + P_{t_f}(r)B_c(r)R_c^{-1}(r)B_c^*(r)P_{t_f}(r)) \times Y_{-B_cK_{t_f}}(r,0)z_0dr
\]
and
\[
Y_{-B_cK_{t_f}}(t,s)z_0 = Y(t,s)z_0 \\
- \int_s^t Y(t,r)B_c(r)R_c^{-1}(r)B_c^*(r)P_{t_f}(r)Y_{-B_cK_{t_f}}(r,s)z_0dr
\]
respectively, have a unique solution \(P_{t_f}(...) \in \mathbb{B}_c^\infty([0,t_f]; \mathcal{L}(H))\). Furthermore, the control
\[
u_{\text{opt}}(t) = -R_c^{-1}(t)B_c^*(t)P_{t_f}(t)z_{\text{adj}}(t)
\]
minimizes the cost (57) and the minimum cost is
\[
J(u_{\text{opt}}(t); z_{\text{adj},0},0) = (P_{t_f}(0)z_{\text{adj},0}, z_{\text{adj},0})_H. \tag{60}
\]

The optimal control problem is now considered on the infinite time interval. Assume that \(A_c(t), B_c(t), \) and \(C_c(t)\) are all linear operators defined for \(t \in \mathbb{R}\) and also that \(B_c(t), C_c(t)\) are bounded uniformly in time. Let \(P_0 \in \mathcal{L}(H), W_c(t) \in C(\mathbb{R}, \mathcal{L}(H))\) and \(R_c(t) \in C(\mathbb{R}, \mathcal{L}(\mathcal{Y}))\) be operators satisfying Assumption 3.1. Both \(W_c(t)\) and \(R_c(t)\) are in addition assumed to be uniformly bounded in time and \(R_c(t)\) is coercive uniformly in time. The cost function becomes
\[
J_\infty(u(t); z_{\text{adj},0}) = \lim_{t_f \to \infty} \int_0^{t_f} \langle (z_{\text{adj}}(r), W_c(r)z_{\text{adj}}(r))_H + \langle u(r), R_c(r)u(r)\rangle_{\mathcal{Y}_0} \rangle dr \tag{61}
\]
subject to the same dynamics (59).

**Theorem A.2** [57, Th. 4.2–4.4]: Assume that for every initial condition \(z_{\text{adj},0} \in H\) in (59), there exists a strongly measurable control input \(u(t)\) such that the cost function (61) is finite. Let \(P_{t_f}(...)\) be the solution to the Riccati (41) and (42) with \(P_0 = 0\). Then, there exists a unique nonnegative and self-adjoint operator \(P_\infty(0) \in \mathcal{L}(H)\) such that, for any \(z_0 \in H\)
\[
\min_u J_\infty(u(t); z_{\text{adj},0},0) = (P_\infty(0)z_{\text{adj},0}, z_{\text{adj},0})_H \tag{62}
\]
\[
\lim_{t_f \to \infty} P_{t_f}(0)z_{\text{adj},0} = P_\infty(0)z_{\text{adj},0} \tag{63}
\]
for any \(z_0 \in H, 0 \leq s \leq t\). Furthermore, for any \(0 \leq s \leq t, P_\infty(s)\) solves
\[
P_\infty(s)z_0 = Y_{-B_cK_\infty}(t,s)P_\infty(t)Y_{-B_cK_\infty}(t,s)z_0 \tag{64}
\]
The optimal control is
\( (68) \)
\( \delta z(t) \)
\( (65) \)

Consider the infinite-time optimal control problem (61) with dynamics (59). Assume that \( A_c(t), W_c(t) \) and \( L(t) \) are exponentially stable in the sense of definition (38). Let \( Y_c(t) \) be the evolution operator generated by \( A_c(t) + L(t)W_c(t) \). Define \( \delta z(t) \) and \( u(t) \) using (66)
\[
\delta z(t) = \int_t^\infty Y_c(t,r) \left( -L(r)W_c(r) \delta z(t) \right) dr + B_c(r)u(t) dr.
\]
From (70)
\[
\sup_{t \in [s,\infty)} \| \delta z(t) \| \mathcal{H} \\
\leq \int_s^\infty \| Y_c(t,r) \left( \sup_{t' \in [s,\infty)} \| L(t') \| \right) \| W_c(t') \delta z(t) \| \mathcal{H} \\
+ \sup_{t' \in [s,\infty)} \| B_c(t') \| \| u(t) \| \mathcal{H} \) dr
\leq \| Y_c(t,s) \| (s,\infty) \sup_{t' \in [s,\infty)} \| L(t') \|
\times \| W_c(t') \delta z(t) \| (s,\infty) \sup_{t' \in [s,\infty)} \| B_c(t') \|
\times \| u(t) \| (s,\infty) \mathcal{H}.
\]
Since \( Y_c(t) \) is exponentially stable, the convergence in (69) into (71) implies that \( \sup_{t \in [s,\infty)} \| \delta z(t) \| \mathcal{H} \) tends to zero. This is equivalent to (67).

Theorem A.4: Consider the infinite-time optimal control problem (61) with dynamics (59). Assume that \( A_c(t), B_c(t) \) are uniformly stabilizable and \( (A_c(t), W_c(t)) \) is uniformly detectable. With the optimal feedback
\[
K_c(t) = R_c(t)B_c(t)P_c(t)
\]
the evolution operator \( Y_{-B_cK_c}(t) \) defined in (65) is exponentially stable.

Proof: The optimal control is
\[
u_{opt}(t) = -K_c(t)z_{adj}(t).
\]
The dynamics of the controlled system are
\[
z_{adj}(t) = Y_{-B_cK_c}(t,s)z_{adj}(s)
\]
with \( u(t) = u_{opt}(t) \) and \( u(t) = u_{t_f}(t) \), respectively. Since \( R_c(t) \) is uniformly lower and upper bounded, \( [57, p. 551] \) implies
\[
J(u_{t_f}, z_{adj}(s), 0) = J_{\infty}(u_{\infty}(t), z_{adj}(s), 0)
\]
where \( u_{\infty}(t) \) and \( z_{adj}(s) \) are the solutions to (66) with \( u(t) = u_{\infty}(t) \) and \( u(t) = u_{t_f}(t) \), respectively. Since \( R_c(t) \) is uniformly lower and upper bounded, \( [57, p. 551] \) implies
\[
J(u_{t_f}, z_{adj}(s), 0) = J_{\infty}(u_{\infty}(t), z_{adj}(s), 0)
\]
and \( Y_{-B_cK_c}(t,s) \) is an evolution operator.
The linear operator \( A_c(t) \) can be reformulated as
\[
A_c(t) = A_c(t) + L(t)W_c(t) - L(t)W_c(t)
\]
where \( L(t) \) is such that the evolution operator generated by \( A_c(t) + L(t)W_c(t) \) is exponentially stable in the sense of definition (38). Let \( Y_c(t,s) \) be the evolution operator generated by \( A_c(t) + L(t)W_c(t) \). Since the dynamics (58) can be represented as
\[
\frac{\partial z_{adj}(t)}{\partial t} = (A_c(t) + L(t)W_c(t) - L(t)W_c(t))z_{adj}(t) + B_c(t)u(t), \quad z_{adj}(s) = z_{adj,0}
\]
z_{adj} can be written
\[
z_{adj}(t) = Y_c(t,s)z_{adj,0} + \int_s^t Y_c(t,r) \left( -L(r)W_c(r)z_{adj}(r) + B_c(r)u(r) \right) dr.
\]
This formulation will be useful in proofs of the following theorems.

Theorem A.3: Consider the infinite-time optimal control problem (61) and finite-time optimal control problem (57) with dynamics (59). Assume that \( (A_c(t), W_c(t)) \) is uniformly detectable. Let \( Y_{-B_cK_c}(t,s) \) be the evolution operator generated by the perturbation of \( A_c(t) \) by \( -B_c(t)R_c(t)B_c(t)P_c(t) \), and similarly \( Y_{-B_cK_c}(t,s) \) is the evolution operator generated by the perturbation of \( A_c(t) \) by \( -B_c(t)R_c(t)B_c(t)P_c(t) \). For each \( z_0 \in \mathcal{H} \)
\[
\lim_{t_f \to \infty} Y_{-B_cK_c}(t,s)z_0 = Y_{-B_cK_c}(t,s)z_0
\]
uniformly in time for \( 0 \leq s < t \leq \infty \).

Proof: Define
\[
u_{\infty}(t) = -R_c(t)B_c(t)P_c(t)z_{adj,\infty}(t)
\]
\[
u_{t_f}(t) = -R_c(t)B_c(t)P_c(t)z_{adj,t_f}(t)
\]
where \( z_{adj,\infty}(t) \) and \( z_{adj,t_f}(t) \) are the solutions to (66) with \( u(t) = u_{\infty}(t) \) and \( u(t) = u_{t_f}(t) \), respectively. Since \( R_c(t) \) is uniformly lower and upper bounded, \( [57, p. 551] \) implies
\[
J(u_{t_f}, z_{adj,0}, 0) = J_{\infty}(u_{\infty}(t), z_{adj,0}, 0)
\]
\[
u_{t_f} \to u_{\infty} \text{ in } L^2([0,\infty); \mathcal{X}_B)
\]
\[
W_c(t)z_{adj}(t) \to W_c(t)z_{adj,\infty}(t) \text{ in } L^2([0,\infty); \mathcal{H}).
\]
Define \( \delta z_{adj}(t) = z_{adj,\infty}(t) - z_{adj,t_f}(t) \) and \( \delta u(t) = u_{\infty}(t) - u_{t_f}(t) \); using (66)
\[
\delta z_{adj}(t) = \int_t^\infty Y_c(t,r) \left( -L(r)W_c(r) \delta z_{adj}(t) \right) + B_c(r)\delta u(t) dr.
\]
exponentially stable. The operator \( A_c(t) - B_c(t) K_{\infty}(t) \) can be written

\[
A_c(t) - B_c(t) K_{\infty}(t) = A_c(t) - \bar{R}_c(t) P_{\infty}(t)
\]

\[
= A_c(t) - L_c(t) W_c^{1/2}(t) + L_c(t) W_c^{1/2}(t) - \bar{R}_c(t) P_{\infty}(t).
\]

Therefore, the evolution operator \( Y_{-B_c K_{\infty}}(t, s) \) generated by \( A_c(t) - B_c(t) K_{\infty}(t) \) can be written

\[
Y_{-B_c K_{\infty}}(t, s) = Y_L(t, s) + \int_s^t Y_L(t, r) \left( L_c(r) W_c^{1/2}(r) - \bar{R}_c(r) P_{\infty}(r) \right) Y_{-B_c K_{\infty}}(r, s) dr.
\]  

(74)

For any initial condition \( v \in \mathcal{H} \)

\[
\| Y_{-B_c K_{\infty}}(t, s) v \|_\mathcal{H} \leq \| Y_L(t, s) v \|_\mathcal{H} + \int_s^t \| Y_L(t, r) \| \| L(r) \| \| W_c^{1/2}(r) Y_{-B_c K_{\infty}}(r, s) v \|_\mathcal{H} dr + \int_s^t \| Y_L(t, r) \| \| \bar{R}_c(r)^{1/2} P_{\infty}(r) \| Y_{-B_c K_{\infty}}(r, s) v \|_\mathcal{H} dr.
\]  

(75)

Let \( \delta_{Y,0}, \alpha_Y > 0 \) be such that \( \| Y_L(t, r) \| \leq \delta_{Y,0} \exp(-\alpha_Y (t - r)) \). Since \( L_c(t) \) and \( \bar{R}(t) \) are uniformly bounded over \([0, t_f]\), for some \( \delta_{Y,0}, \delta_{Y,1}, \delta_{Y,2} > 0 \)

\[
\| Y_{-B_c K_{\infty}}(t, s) v \|_\mathcal{H} \leq \delta_{Y,0} \exp(-\alpha_Y (t - s)) \| v \|_\mathcal{H} + \delta_{Y,1} \int_s^t \exp(-\alpha_Y (t - r)) \| W_c^{1/2}(r) Y_{-B_c K_{\infty}}(r, s) v \|_\mathcal{H} dr + \delta_{Y,2} \int_s^t \exp(-\alpha_Y (t - r)) \| \bar{R}_c(r)^{1/2} P_{\infty}(r) \| Y_{-B_c K_{\infty}}(r, s) v \|_\mathcal{H} dr.
\]  

(76)

Since \( W_c^{1/2}(r) Y_{-B_c K_{\infty}}(r, s) v \in \mathbb{L}^2([0, \infty); \mathcal{H}) \) and \( \bar{R}_c(r)^{1/2} P_{\infty}(r) Y_{-B_c K_{\infty}}(r, s) v \in \mathbb{L}^2([0, \infty); \mathcal{H}) \) (73), the convolution product

\[
\| Y_{-B_c K_{\infty}}(t, s) v \|_{\mathbb{L}^2([0, \infty); \mathcal{H})} \leq \delta_{Y,3}
\]  

(77)

where \( \delta_{Y,3} > 0 \) is independent of \( t_f > 0 \). (See for example [51, Lemma A.6.6].) Furthermore, \( Y_{-B_c K_{\infty}} \) is a continuous evolution operator by [41, Th 1.1] and has an exponential growth bound. Therefore, by Datko’s theorem for evolution operators, [58, Th 1.1], \( Y_{-B_c K_{\infty}}(t, s) \) is exponentially stable. \( \square \)

Theorem A.5: Assume that \( (A_c(t), B_c(t)) \) is uniformly stabilizable and \( (A_c(t), W_c^{1/2}(t)) \) is uniformly detectable. Then, the solution \( P_{\infty}(s) \) to the Riccati (64) is unique and yields the optimal cost defined by (57).

Proof: Let \( u(t) \) be any admissible control input strongly measurable control input such that (61) is finite. The cost function (61) is bounded above and thus for some \( \delta_u > 0 \)

\[
\lim_{t_f \to \infty} \int_s^{t_f} (u(r), R_c(r) u(r)) x_0 dr \leq \delta_u
\]

\[
\lim_{t_f \to \infty} \int_s^{t_f} (z_{adj}(r), W_c(r) z_{adj}(r)) x_0 dr \leq \delta_u.
\]  

(78)

Since \( R_c(t) \) is uniformly coercive

\[
u(t) \in \mathbb{L}^2([s, \infty); X_B).
\]  

(79)

Using (66) and letting \( M_L, \alpha_L \) be such that \( \| Y_L(t, s) \| \leq M_L \exp(-\alpha_L (t - s)) \)

\[
\| z_{adj}(t) \|_\mathcal{H} \leq \| Y_L(t, s) \| \| z_{adj,0} \|_\mathcal{H} + \int_s^t \| Y_L(t, r) \| \| L(r) \| \| W_c^{1/2}(r) z_{adj}(r) \|_\mathcal{H} dr + \int_s^t \| Y_L(t, r) \| \| B_c(r) \| \| u(r) \|_U dr \leq M_L \| z_{adj,0} \|_\mathcal{H} \exp(-\alpha_L (t - s)) + M_L \sup_r \| L(r) \| \| W_c^{1/2}(r) z_{adj}(r) \|_\mathcal{H} dr + M_L \sup_r \| B_c(r) \| \times \int_s^t \exp(-\alpha_L (t - r)) \| u(r) \|_U dr.
\]  

(80)

The convolution product (80) and (78) implies that

\[
z_{adj}(t) \in \mathbb{L}^2([s, \infty); \mathcal{H}).
\]  

(81)

(See, for example, [51, Lemma A.6.5].)

It will now be shown that \( z_{adj}(t) \to 0 \) using the technique in the proof of [57, Lemma 4.1]. Due to (79) and (81), for any \( n \) there exists \( t_n > 0 \) such that

\[
\int_{t_n}^{\infty} \| z_{adj}(r) \|_\mathcal{H}^2 dr < \frac{1}{n^2}, \int_{t_n}^{\infty} \| u(r) \|_U^2 dr < \frac{1}{n^2}.
\]

Define \( S_t = \{ t | t > t_n \& \| z_{adj}(t) \|_\mathcal{H} \geq 1/n^2 \} \). Note that

\[
\int_{S_t} \| z_{adj}(r) \|_\mathcal{H}^2 dr < \int_{S_t} \| u(r) \|_U^2 dr < \frac{1}{n^2}
\]

and so \( \int_{S_t} \| u \|_U^2 < 1/n^2 \); write \( \delta_n = \int_{S_t} \| u \|_U^2 dr \). Thus, for \( t \geq t_n + 1/n^2 \), there exists \( s \) between \( t_n \) and \( t \) such that \( |t - s| < \delta_n \) and \( \| z_{adj}(s) \| < 1/n \). From (59)

\[
\| z_{adj}(t) \|_\mathcal{H}^2 \leq \delta_{Y,0} \exp(\alpha_{Y,0} (t - s)) \frac{1}{n} + \delta_{Y,0} \int_s^t \exp(\alpha_{Y,0} (r - s)) \| B_c(r) \| \| u(r) \|_U dr \leq \delta_{Y,0} \exp(\alpha_{Y,0} \delta_n) \left( 1/n + \sup_r \| B_c(r) \| \left( t - s \right)^{1/2} \left( \int_s^\infty \| u(r) \|_U^2 dr \right)^{1/2} \right) \leq \delta_{Y,0} \exp(\alpha_{Y,0} \delta_n) \left( 1/n + \sup_r \| B_c(r) \| 1/n^3 \right).
\]
Therefore
\[ \lim_{t \to \infty} \| z_{adj}(t) \| = 0. \]
Theorem 4.6 in [57] then implies that there is at most one solution to the Riccati integral (64) on the infinite interval and hence \( P_\infty(s) \) is the optimal cost.

**Theorem A.6:** The solution \( P_{t_f}(0) \) to (41) can be uniformly bounded independent of \( t_f > 0 \). With the optimal feedback
\[ K_{t_f}(t) = R_{c}^{-1}(t)B_{c}^r(t)P_{t_f}(t) \]
the evolution operator \( Y_{-B_{c}K_{t_f}}(t, s) \) defined in (42) is uniformly bounded independent of \( t_f \).

**Proof:** Since \( P_\infty(0) \) yields the optimal cost to the infinite-time problem, for all \( t_f > 0 \) and \( z_{adj,0} \in \mathcal{H} \)
\[ (P_{t_f}(0)z_{adj,0}, z_{adj,0})_\mathcal{H} \leq (P_\infty(0)z_{adj,0}, z_{adj,0})_\mathcal{H}. \]
Thus, there exists \( \delta_p > 0 \) such that for all \( t_f > 0 \)
\[ \| P_{t_f}(0) \| \leq \delta_p. \]
Now, from optimality of \( u_{opt}(t) = -K_{t_f}(t)z_{adj}(t) \), and definition of the cost (57) it can be concluded that for some \( \delta_p > 0 \)
\[ \int_0^{t_f} \| W^1/2(r)z_{adj}(r) \|^2_dr \leq \delta_p \| z_{adj,0} \|^2_\mathcal{H}, \]
\[ \int_0^{t_f} \| \tilde{R}_{c}^1/2(r)P_{t_f}(r)z_{adj}(r) \|^2_rdr \leq \delta_p \| z_{adj,0} \|^2_\mathcal{H}. \]
and for \( u(t) = u_{opt}(t) \), (66) leads to, letting \( M_L, \alpha_L \) be such that
\[ \| Y_L(t, s) \| \leq M_L \exp(-\alpha_L(t - s)) \]
\[ \| z_{adj}(t) \|_\mathcal{H} \leq \| Y_L(t, s) \| \| z_{adj,0} \|_\mathcal{H} \]
\[ + \int_s^t \| Y_L(t, r) \| \left( \sup_{r'} \| L(r') \| \| W^1/2(r)z_{adj}(r) \|_\mathcal{H} \right) \]
\[ + \sup_{r'} \| \tilde{R}_{c}^1/2(r') \| \| \tilde{R}_{c}^1/2P_{t_f}(r)z_{adj}(r) \|_r \| t \|_r \]
\[ \leq M_L \exp(-\alpha_L(t - s)) \| z_{adj,0} \|_\mathcal{H} \]
\[ + \frac{M_L}{\sqrt{2\alpha}} \sup_{r'} \| L(r') \| \| W^1/2(t)z_{adj}(t) \|_{L^2([s, \infty]; \mathcal{H})} \]
\[ + \frac{M_L}{\sqrt{2\alpha}} \sup_{r'} \| \tilde{R}_{c}^1/2(r') \| \| \tilde{R}_{c}^1/2P_{t_f}(r)z_{adj}(r) \|_{L^2([s, \infty]; \mathcal{H})} \]
\[ \leq M_L \left( \exp(-\alpha_L(t - s)) + \frac{2\delta_p}{\sqrt{2\alpha}} \right) \| L(r') \| \]
\[ + \sup_{r'} \| \tilde{R}_{c}^1/2(r') \| \| z_{adj,0} \|_H. \]
Thus, \( \| Y_{-B_{c}K_{t_f}}(t, s) \| \) is a bounded operator independent of \( t_f > 0. \)

**ACKNOWLEDGMENT**

The authors would like to thank Shuxia Tang for her comments on early version of this article.

**REFERENCES**

[1] A. Bensoussan, *Filtrage optimal des systèmes linéaires*. Paris, France: Dunod, 1971.
[2] R. F. Curtain, “A survey of infinite-dimensional filtering,” *SIAM Rev.*, vol. 17, no. 3, pp. 395–411, 1975.
[3] A. Bensoussan, “Some remarks on linear filtering theory for infinite dimensional systems,” in *Directions in Mathematical Systems Theory and Optimization* (Lecture Notes in Control and Information Sciences Series), A. Rantzer and C. I. Byrnes, Eds., vol. 286. Berlin, Germany: Springer-Verlag, 2003, pp. 27–39.
[4] A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter, *Representation and Control of Infinite Dimensional Systems*. Berlin, Germany: Springer, 2007.
[5] X. Wu, B. Jacob, and H. Elberm, “Optimal control and observation formulations for time-varying systems on infinite-time horizon,” *SIAM J. Control Optim.*, vol. 54, no. 1, pp. 291–316, 2015.
[6] A. Smyshlyaev and M. Krstić, “Backstepping observers for a class of parabolic PDE’s,” *Syst. Control Lett.*, vol. 54, no. 7, pp. 613–625, 2005.
[7] R. Miranda, I. Chairez, and J. Moreno, “Observer design for a class of parabolic PDE via sliding modes and backstepping,” in *Proc. IEEE 11th Int. Workshop Variable Struct. Syst.*, 2010, pp. 215–220.
[8] R. Curtain, “Finite-dimensional compensator design for parabolic distributed systems with point sensors and boundary input,” *IEEE Trans. Autom. Control*, vol. 27, no. 1, pp. 98–104, Feb. 1982.
[9] M. A. Demetriou, “Natural second-order observers for second-order distributed parameter systems,” *Syst. Control Lett.*, vol. 51, no. 3, pp. 225–234, 2004.
[10] S.-X. Tang, L. Camacho-Solorio, Y. Wang, and M. Krstić, “State-of-charge estimation from a thermal–electrochemical model of lithium-ion batteries,” *Automatica*, vol. 83, pp. 206–219, 2017.
[11] S. Tang and M. Krstić, “Observer design for an IPDE with time-dependent coefficients,” in *Proc. Amer. Control Conf.*, 2016, pp. 1655–1660.
[12] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Hoboken, NJ, USA: Wiley, 2006.
[13] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice Using MATLAB*. Hoboken, NJ, USA: Wiley, 2011.
[14] K. Reif and R. Unbehauen, “The extended Kalman filter as an exponential observer for nonlinear systems,” *IEEE Trans. Signal Process.*, vol. 47, no. 8, pp. 2324–2328, Aug. 1999.
[15] K. Reif, S. Günther, E. Yaz, and R. Unbehauen, “Stochastic stability of the discrete-time extended Kalman filter,” *IEEE Trans. Autom. Control*, vol. 44, no. 4, pp. 714–728, Apr. 1999.
[16] K. Reif, S. Günther, E. Yaz, and R. Unbehauen, “Stochastic stability of the continuous-time extended Kalman filter,” *IEEE Proc.-Control Theory Appl.*, vol. 147, pp. 45–52, 2000.
[17] X. Kai, C. Wei, and L. Liu, “Robust extended Kalman filtering for nonlinear systems with stochastic uncertainties,” *IEEE Trans. Syst., Man Cybern., Part A: Syst. Hum.*, vol. 40, no. 2, pp. 399–405, Mar. 2010.
[18] X. Kai, L. Liangdong, and L. Yiu, “Robust extended Kalman filtering for nonlinear systems with multiplicative noises,” *Optimal Control Appl. Methods*, vol. 32, no. 1, pp. 47–63, 2011.
[19] G. A. Einicke and L. B. White, “Robust extended Kalman filtering,” *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2596–2599, Sep. 1999.
[20] S. Elizabeth and R. Jothilikashmi, “Convergence analysis of extended Kalman filter in a noisy environment through difference equations,” *Int. J. Appl. Math.*, vol. 91, no. 1, pp. 33–41, 2014.
[21] F. Alonge, T. Cangemi, F. D’Ippolito, A. Fagiolini, and A. Sferlazza, “Convergence analysis of extended Kalman filter for sensorless control of induction motor,” *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2341–2352, Apr. 2015.
[22] M. Boutaye, H. Rafaralahy, and M. Darouach, “Convergence analysis of the extended Kalman filter used as an observer for nonlinear deterministic discrete-time systems,” *IEEE Trans. Autom. Control*, vol. 42, no. 4, pp. 581–586, Apr. 1997.
[23] J. Baras, A. Bensoussan, and M. James, “Dynamic observers as asymptotic limits of recursive filters: Special cases,” *SIAM J. Appl. Math.*, vol. 48, no. 5, pp. 1147–1158, 1988.
[24] K. Reif, F. Somemmann, and R. Unbehauen, “An EKF-based nonlinear observer with a prescribed degree of stability,” *Automatica*, vol. 34, no. 9, pp. 1119–1123, 1998.
[25] J. H. Ahrens and H. K. Khalil, “Closed-loop behavior of a class of nonlinear systems under EKF-based control,” *IEEE Trans. Autom. Control*, vol. 52, no. 3, pp. 536–540, Mar. 2007.
