Observability Based Techniques to Analyze and Design User-Interfaces

Situation-Awareness and Displayed Information

by

Neda Eskandari Naddaf

B.Sc., Sharif University of Technology, 2006
M.Sc., Sharif University of Technolog, 2008

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2015

Abstract

For continuous-time LTI systems under human control and under shared control, this thesis studies techniques to determine whether or not a given user-interface provides the information required to accomplish a certain task. It is well known that attaining Situation Awareness (SA) is essential to the safe operation of the systems involving human-automation interaction. Hence, through two different approaches, the work in this thesis evaluates and designs user-interfaces based on the satisfaction of SA requirements by the user.

In the first approach, observability-based conditions under which a user-interface provides the user with adequate information to accomplish a given task are identified. The user is considered to be a special type of observer, with capabilities corresponding to different levels of knowledge regarding the input and output derivatives. Through this approach, the "user-observable/user-predictable" subspaces for systems under shared control are defined and formulated. In addition, state estimation is considered to incorporate a processing delay. Hence, the "delay-incorporating user-observable/user-predictable" subspaces are formulated and are compared with the space spanned by the combination of the states which create the task. If the task subspace does not lie in the relevant space, then the user-interface is incorrect, meaning that the user cannot accomplish the desired task with the given user-interface.

In the second approach, to determine the required information to be displayed, a model of attaining SA for the users is proposed. In this model, the user is modeled as an extended delayed functional estimator. Then, the information needed for such an estimator to make correct estimations as well as the desired expansion of the functional of the states to let the user precisely reconstruct and accurately predict the desired task is determined. Additionally, it is considered that in practice, to attain the situation awareness, the estimation of the task states does not necessarily need to be precise but can be bounded within certain margins. Hence, the model of the user attaining SA is also modified as a "bounded-error delayed functional observation/prediction". Such an observer/predictor has to exist for a system with a given user-interface, otherwise, the safety of the operation may be compromised.

Preface

My contributions in this thesis resulted in two conference papers and five journal articles.

- A version of Chapter 3 is published in:

  N. Eskandari, and M. Oishi, "Computing observable and predictable subspaces to evaluate user-interfaces of LTI systems under shared control," in *Proceedings of the IEEE conference on Systems, Man and Cybernetics*, Alaska, USA, October 2011, pp. 2803–2808. This paper received the Best Student Paper Award at this conference.

  R. Cortez, D. Tolic, I. Palunko, **N. Eskandari**, R. Fierro, M. Oishi, and J. Wood, A hybrid framework for user-guided prioritized search and adaptive tracking of maneuvering targets via cooperative UAVs, In *Advances in Intelligent and Autonomous Aerospace Systems*, in the *Progress in Aeronautics and Astronautics Series* American Institute of Aeronautics and Astronautics, Reston, VA, vol. 241, 2012.

- The results from Chapter 4 are published in:

  N. Eskandari, G. A. Dumont, and Z. J. Wang, "Delay-incorporating observability and predictability analysis of safety-critical continuous-time systems," *IET Control Theory and Applications*, vol. 9, pp. 1692–1699(7), July 2015.

- The results from Chapter 5 are published in:

  N. Eskandari, G. A. Dumont, and Z. J. Wang, "An observer/predictor based model of the user for attaining situation awareness," *IEEE Transactions on Human-Machine Systems*, 2015.

- The results from Chapter 6 were submitted as:

  Neda Eskandari, Guy Dumont, and Z. Jane Wang. "Bounded-error delayed functional observer/predictor: Existence and design".

- The results from Appendix A are published in:

  N. Eskandari, M. M. Oishi, and Z. J. Wang, "Observability analysis of continuous-time LTI systems with limited derivative data," *Asian Journal of Control*, vol. 16, no. 2, pp. 623–627, 2014.

- The results from Appendix B are published in:

  N. Eskandari, Z. J. Wang, and G. Dumont, "Modeling the user as an observer to determine display information requirements," in *In Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, Manchester, UK, October 2013, pp. 267–272.

  N. Eskandari, Z. J. Wang, and G. A. Dumont, "On the existence and design of functional observers for LTI systems, with application to user modeling," *Asian Journal of Control*, 2014.

I, hereby, declare that I am the first author of this thesis. I conducted the literature survey, identified the problem to solve, formalized the theoretical solution and performed the analysis, and finally implemented all the simulations. In addition, I am responsible for writing all of the drafts of this thesis and the resulted papers. My supervisors Professor Guy A. Dumont and Z. Jane Wang guided my research and provided me with complete support including validating the methodologies and editing all the manuscripts co-authored by them. Part of the thesis is the result of research collaboration with an additional contributor, Dr. Meeko Oishi for her technical and editing feedbacks.

Table of Contents

Appendices

List of Tables

List of Figures

List of Abbreviations and Symbols

Acronyms

DOA	Depth of Anesthesia
GPS	Global Positioning System
LTI	Linear Time Invariant
OCM	Optimal Control model
PD	Pharmacodynamic
PK	Pharmacokinetic
SA	Situation Awareness
SGA	Situation Goal Output
THS	Trimmable Horizontal stabilizer

Notations and Operators

Many notations are introduced in the body of this thesis with their description provided in the respective sections. Following, we provide the main important notations and operators that are used more frequently throughout this document:

$\mathcal{O}_{\mathcal{H}}$	User-observable subspace
$\mathcal{P}_{\mathcal{H}}$	User-predictable subspace
$\mathcal{O}^*_{\mathcal{H}}$	Delay-incorporating user-observable subspace
$\mathcal{P}^*_{\mathcal{H}}$	Delay-incorporating user-predictable subspace

$\mathcal{O}_{\mathrm{UI}}$	Unknown-input observability subspace
$\mathcal{O}_{H,\tau_1}$	Delayed observable space
$\mathcal{T}$	Task space
$\mathcal{C}$	Controllability subspace
$\mathcal{O}$	Observability subspace
λ	Available number of derivatives of input
γ	Available number of derivatives of output
τ	Prediction horizon
τ_1	Processing delay
$\mathcal{R}(.)$	Column space
$\mathcal{N}(.)$	Null space
$(.)^{\perp}$	Orthogonal complement
$rank(.)$	Rank of a matrix
$\mathbb{R}$	Real numbers
$\mathbf{I}_N$	$N \times N$ identity matrix
$\mathrm{diag}\{\cdot\}$	(Block) Diagonal matrix
$\oplus$	Subspace summation
$\cap$	Intersection

Acknowledgments

I am using this opportunity to express my gratitude to my supervisors Professor Guy A. Dumont and Professor Z. Jane Wang without whom this research would not be possible. Their continuous understanding, support, and patience helped me to overcome a series of difficulties and conduct the presented research. I would also like to thank my supervisors for all the technical guidance that they provided me with.

I also wish to deeply thank my loved ones who have supported me throughout the entire process. Many thanks to my dear husband, Hossein, who has been incredibly supportive of me and my work since the very first day of our marriage, more than eleven years ago! His selfless patience and his faith in me throughout the difficult phases of my Ph.D. was one of the main reasons that this work could come to existence. Words cannot express how grateful I am to my lovely parents for all of the sacrifices that they have made for my success. Their consistent support and also prayers for me was what sustained me thus far. I would also like to thank my brother and my sister-in-law who have always been available for me during this Ph.D..

I am grateful to my friends and colleagues for helping me throughout this process.

Chapter 1

Introduction

1.1 Motivation

In complex cyber-physical systems, correct human interaction with the system is key for effective operation. Oftentimes, such systems are very large and simple intuition is not enough to determine whether a *user-interface*, a device through which the human applies a control input and through which the system provides the human with information about the output, displays the required information. Not only the representation of information, but also the information content of the display, play important role in having a good interface. While too much information can overwhelm the user, presenting too little information can result in non-determinism in the user-interface from the user's point of view [4]. Mathematical tools and methods to determine whether a given user-interface allows accomplishment of a feasible task can help prevent non-determinism and other inconsistencies that could arise through incorrect user-interface design.

Consider an example about 1994 Nagoya Airbus A300 incident [5]. For this flight, an automatic go-around maneuver was inadvertently triggered during a manual approach to landing. While the flight crew applied a pitch-down command to achieve their desired path, automation applied a pitch-up command to reach a reference altitude rapidly. Unaware of the effect of their input, cabin crew continued to descend along the glideslope. The automation adjusted the trimmable horizontal stabilizers (THS) to make the aircraft ascend, while the flight crew acted to counter the automation through the elevators. Meanwhile, the THS gradually moved from $-5.3°$ to $-12.3°$, producing an orientation very close to the nose-up limit. The flight crew eventually disengaged the autopilot and decided to abort the landing. A pitch-up command to the elevators, in combination with the THS orientation, generated a stall condition which resulted in a crash [6]. Although it might not seem obvious that under given situation, the task of landing was not a safe task, mathematical tools can be developed to help the designers come up with such a conclusion.

Researchers have previously identified potential sources of mode confusion [7] in discrete user-interfaces for aircraft flight management systems modeled as discrete-event systems [4, 8–10] and as discrete abstractions of hybrid systems [6, 11]. Methods based on model-checking evaluate a discrete event system that represents the underlying system dynamics and user-interface for deadlock and other problematic states. Methods based on composition of finite state machines compare the evolution of trajectories in a discrete event system that represents the underlying system, to another discrete event system that maps the states to a known output, essentially providing a simplified representation of the information on the user-interface that evolves according to the system dynamics [12]. However, while continuous components of the interface are quite common (e.g., altimeter, speedometer, others), little work has been done on evaluating the correctness of user-interfaces with continuous components that have not been abstracted to discrete states [4, 13].

As interfaces and the underlying systems become more complex, information beyond what is contained in the interface may not be accessible. We aim to identify tools that assess the correctness of a user-interface for a given task, an especially relevant problem in systems for which intuition and simulation may not be enough to assure that an interface is effective. While in many systems, such as a human-driven car, the user has access to information beyond what is simply contained in the interface, we focus here solely on information contained in the interface, e.g., a remote operator controlling a fleet of UAVs or a pilot performing a task in high altitude.

Using the mathematical tools that we develop, we can demonstrate that for the aircraft in the above example under shared control, the display did not provide the user with sufficient information to perform the task of landing safely. More importantly, by making further investigations we can show that even with further information in the display e.g. about the automation desired trajectory, accomplishing landing is not a safe task to be done under shared control. So given these results, we can suggest either providing the user with a predicted information about the process of landing or disengaging the autopilot during the phase of landing.

1.2 Situation awareness and display design

Despite successful efforts to increase the autonomy of systems and devices used daily, many systems function under the shared control of the human and the computer. Indeed many fully automated systems need to be supervised and monitored by a human operator. Because correct interaction between the user and the plant is crucial, it is essential for the user to understand what the system is doing and what it is intending to do in a near future. For this, the system should provide the required information for the user to achieve complete Situation Awareness (SA), i.e. to ”keep the operator tightly coupled to the dynamics of the environment” [14].

Endsley formally defines SA as ”the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future” [3]. The schematic framework for SA that was suggested by Endsley in [3] is shown in Figure 1.1.

From Figure 1.1 and the above definition, three components of SA are as follows [3, 15]:

- Level 1: Perception. Level 1 of SA is about being aware of the existence of the elements in the environment, the dynamics of the system, and more. Lack of correct perception can increase the chances of wrong understanding about the situation.

- Level 2: Comprehension. Based on the particular goal of the user, each part of the perceived information in Level 1 can be of certain importance. Comprehension of information is about integrating and

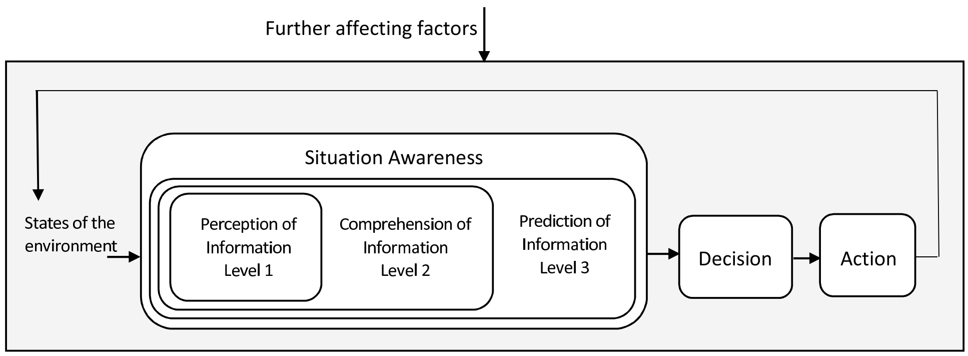


Figure 1.1: Model of SA, from [3]. As is shown in the figure, SA which consists of three levels of processing the information is necessary for the user to make correct decisions on the control action.

synthesizing this information to determine the significance of each portion of the information and to attain further and possibly high level understanding about the environment.

- Level 3: Prediction. The highest stage of SA is projection which is about using all information obtained in Level 1 and 2 to make predictions about the future states of the system and the elements of the environment.

As is clear from Figure 1.1, having SA about the desired task is necessary for the user to be able to make correct decisions and finally to choose the reasonable and safe control action. Therefore, it is of interest to develop techniques to evaluate and determine the minimum information required for the user to achieve SA.

The tools that we develop for analyzing available displays as well as designing better displays are based on the theory of SA. For instance, for the aircraft incident described in Section 1.1, we will later show that the available display of A300 did not provide the pilot with adequate information to achieve SA about the task of landing which finally led to a crash.

1.3 Research objectives

As has already been discussed, it is well-known from the literature that attaining SA is key for the user to have correct interaction with the automation [3, 16]. Not having SA may lead to wrong decision-making and possibly a faulty control action by the operator. Attaining SA necessitates three stages of processing the information, i) perception of the information, ii) comprehension of the information, and iii) projection or prediction of the information [3, 16].

With the purpose of analyzing and modeling the process of attaining SA and its relationship with displayed information, this thesis considers the cases for which the only source of information for the user is the display. Therefore, the display has to be designed to allow the user to perceive, comprehend, and predict the desired information to attain correct SA. More specifically, the display design requires a careful selection and clear presentation of the information for the user to perform those tasks properly.

To achieve the main objective of this thesis which is determining the required displayed information, observability-based conditions under which a user-interface provides the user with adequate information to accomplish a given task are identified. In addition, a model of attaining SA for the users is proposed.

1.4 Related work

Cognitive modeling has long been a topic of interest for human-factors researchers, engineers, clinical psychologists, and scientists in any other field which is creating a link between human behavior and mathematical concepts [17–19]. As opposed to conceptual models which discuss and evaluate the specifications of the system and the interactions between the user and the plant in verbal terms [20], cognitive models use analytical tools and techniques to provide more precise and valid results and predictions of the behavior of such systems [21]. The main difference between a cognitive model and a statistical or a generic mathematical model is that in a cognitive model certain specifications of the human and cognitive parameters are considered to customize the analytical model and to address the behavior of the human.

Although, in general, mostly the statistical based models of human are considered and named as cognitive models [22–24], we also can consider the models of human suggested in other types of papers as cognitive models. For instance, over the past few decades, much work has been done on modeling the human operator as a controller [25–28]. Each of these models includes various parts representing different specifications of the human, such as, neuromuscular term, proprioceptive feedback, and delay and in each paper the results are compared with the real-life results. In more recent papers, uncertainties were also incorporated in the human models [29, 30].

Because SA is essential for the user to properly accomplish a task, it is necessary to either 1) analyze the processing of the information by the user to evaluate whether and/or how the user is capable of attaining SA about the important states of the system or 2) to extend the human model to also include obtaining SA before acting as a controller.

Several researchers have investigated conceptual models of human information processing for attaining SA in various applications [31–33]. However none of these papers quantified this process with a rigorous analytical model.

To our knowledge, the optimal control model (OCM) [34, 35] was the first model to mathematically capture the ability of the user as an estimator in addition to a controller. In general, in the OCM, the user estimates the states of the system and acts on the system based on these estimations. This model states that the human attains delayed noisy information from the display then makes further estimations about the states of the system. The estimator block in the OCM is considered to be a Kalman Filter. A predictor block is also included in this model which simply compensates for the inherent delay of the user. This model had aimed at capturing some limitations of the human, including the time delay, neuro-motor dynamics, and some controller specifications. However, given that the time delay is considered to be canceled out by the predictor and no further effort is made to customize the Kalman filter based on the limitation of the user, the estimation part of this model is more of a mathematical model than a cognitive model.

In addition to the above models, other researchers also modeled the human as an observer-based fault detector with the focus on modeling the decision-making process [36, 37].

1.4.1 Observability and predictability

Observability analysis and observer design are our main tools through this thesis.

The concept of observability was first introduced by Luenberger [38, 39] for single and multi-variable systems. Later, through using a geometrical approach based on the controlled and conditioned invariant spaces, Basile and Marro [40] could formulate the observability subspace for systems with unknown input

By taking derivatives from the output equation in

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx \end{aligned} \tag{1.1}$$

where $\mathcal{R}(.)$ shows the column space and based on definition, we have

Definition 1. *The least $(A, \mathcal{F})$-conditioned invariant containing a space $\mathcal{X}$ is the subspace*

$$\mathcal{J}_m = \mathcal{Y}_{n-1}$$

where $\mathcal{Y}_{n-1}$ is defined by a recursive relationship

$$\mathcal{Y}_0 = \mathcal{X}, \qquad \mathcal{Y}_i = \mathcal{X} \oplus A(\mathcal{Y}_{i-1} \cap \mathcal{F}), \qquad i = 1, \cdots, n-1.$$

In Chapters 3 and 4, we draw mainly on control and observability concepts by Kalman [41] and Luenberger[38, 39], as well as the work by Basile and Marro [40] in unknown-input observability to obtain the user-observable subspace – that is, a subspace which is observable to the user with additional limitations.

1.4.2 Observer and predictor design

In Chapters 5 and 6, we model the user as a functional observer under a set of given assumptions about the actual behavior of the user.

For many applications, such as implementing a control law, monitoring the behavior of an automated system, or for fault detection, one needs to have access to all or a portion of the state space. If the desired portion of the state space is not directly measured, an observer is required to estimate the desired states based on the measured states. Over the past few decades, numerous research papers have been published on observing the states of linear systems. Luenberger initially came up with the ideas of designing full-state observers to estimate all states of a deterministic linear system and reduced-order observers to reconstruct the unmeasured states of a deterministic linear system [38, 39]. In many applications, however, there is no need for reconstructing all unmeasured states, since only a specific portion of the state space is important to the users/designers (e.g., for state feedback control). Therefore, to reduce the cost of reconstructing unnecessary states, functional observers were introduced.

Functional observers are a specific type of observers which can reconstruct a desired functional without reconstructing all unmeasured states. The first functional observer, which was capable of reconstructing a scalar linear functional of the states of a linear system, was presented in [39]. This technique was followed by other efforts [42, 43] and was later extended for observing multi-functionals of LTI systems [42, 44, 45].

While designing observers for the reconstruction of linear functionals remains a popular topic, researchers are also interested in investigating methods for reducing the order of the functional observers [46] and for evaluating the minimum required order [47, 48]. An excellent piece of work [49] presents a necessary and sufficient condition for the existence of a "same order functional observer" (i.e., a functional observer with the same order as the desired functional). Later [50], [51], and [52] extended [49] to evaluate the existence of "potentially higher order observers" that can reconstruct the desired functional in cases that the observer in [49] does not exist. Another approach for assessing the existence of functional observers is the eigenspace analysis [53].

Functional observers are of great importance, not only for systems with known input, but also for systems with unknown input. Researchers have made efforts to design functional observers for systems with unknown input [54–57]. As in his work on systems with known input and no uncertainty [49], in [58] Darouach defined necessary and sufficient conditions for the existence of a "same order functional observer" for systems containing unknown inputs. Further, by modifying [58], more general existence conditions for unknown input functional observers were reported in [59].

Researchers studied the existence of functional observers for the reconstruction of desired functionals for different types of systems, including linear descriptor systems [52, 60], time delay systems [61–65], and two dimensional systems [66]. Also functional observers are designed for the purposes such as fault detection [67, 68]. There are also works on the existence and design of common/simultaneous functional observers [69, 70]. The references mentioned here are only a portion of the vast body of literature on functional observers for linear systems.

In most of the above papers on the existence and design of functional observers, the overall procedure for designing the functional observer for an LTI system is as follows:

- Assume to have a parametrized observer in an LTI form. Based on the availability of the input, the input will be introduced to the dynamics of the observer.
- Formulate the error of estimation to be the difference between the estimated functional and the actual functional of the states.
- Find the dynamics of the error by taking derivative from the estimation error we formulated before.
- Formulating the error dynamics as $\dot{e} = C_1 e + C_2$ where C_1 and C_2 are functions of the observer matrices.

- To make the error asymptotically approach zero, C_1 should be designed to be stable and C_2 have to be zero.

The effect of the availability of the delayed output on the size of the desired functional has also been evaluated [71–73]. A rather comprehensive discussion of functional observers for systems with known and unknown inputs is available in [74].

1.5 Summary of contributions

Chapters 3 through 6 cover techniques for the analysis and syntheses of user interfaces to satisfy the requirements of SA theory.

In Chapter 3, we identify observability-based conditions under which user-interfaces provide the user with information to accomplish a given task, formulated as a subset of the state-space. We introduce the "user-observable subspace" and the "user-predictable subspace" which are subspaces based on observability and predictability requirements of the user and limitations of the user regarding input signals.

The main contribution of Chapter 4 is to modify the results of Chapter 3 by incorporating the delay of estimation. Therefore, the notions of delay-incorporating user-observable and delay-incorporating user-predictable spaces are defined and are formulated in this chapter. The delay that we consider throughout this thesis is the information processing delay of the user. On the other hand, our system of interest is delay free.

In Chapter 5, we model SA with specific consideration of the capabilities of the user. We, thus, evaluate the existence of the novel delayed observers/predictors and then design them. We, also, suggest a technique for determining the required displayed information.

The work in Chapter 6 comes up with an estimator model that makes bounded estimations of the desired states of a deterministic LTI system. Therefore, a novel technique to check the boundedness of the estimation and the prediction errors of a desired functional, with an estimator which dynamics is delayed, is suggested.

In the appendices, the effect of providing higher derivatives of the input and/or the output signal on the observability subspaces and on the order and the structure of the functional observer is investigated.

Chapter 2

Problem definition

Incorrect human-automation interaction can be due to reasons such as non-deterministic plant behavior or the user having inadequate understanding of the plant dynamics. In addition, inadequate and improper display of information could be hazardous to the safety of the system.

2.1 Problem statement

Consider a user interacting with an automated system through the user-interface. Building a reliable interaction between this user and the automated plant requires providing the user with necessary information about the states of the plant and to set up a proper task for him/her to accomplish. Through processing the available information, the user can achieve situation awareness and then use it to accomplish a desired set of tasks.

Various researchers have examined human information processing [3, 75–77]. Parasuraman et. al. [76] introduced a four-stage model of human information processing which includes information acquisition, information analysis, decision and action selection, and action implementation. According to the situation awareness model of Endsley [3], the user first attains situation awareness. Decision making and action implementation are two stages that follow situation awareness. Sherry's *SGA* model of human information processing, which was later used by Sherry et.al. [77] to demonstrate one of the reasons for users' confusion about the behavior of automation, also consists of attaining information about the situation, re-scheduling the goals, and finally action.

Essentially, three common fundamental stages of human information processing have been pointed out by the above mentioned researchers, including

i) understanding the situation, ii) decision making, and iii) action implementation. In this work, we focus on the first stage of information processing – that is, understanding the situation. Our main goal here is to evaluate the displayed information and obtain the required information which has to be displayed so that the user can accomplish a desired task.

Analyzing the available information has been discussed in detail by Endsley and it was specifically termed as Situation Awareness (SA) [3]. Attaining SA which necessitates perception, comprehension, and prediction of the information, can be done using *working memory or long term memory.* As stated by Cowan [78], "long-term memory is a vast store of knowledge and a record of prior events" and the working-memory is the memory which "is used to plan and carry out the behavior". Other researchers also mentioned that understanding and predicting the required information is vital for having an effective human-automation interaction and that problems with predictability can result in false expectations [7, 79].

Factors such as learning and attention can clearly affect the perception of the information (the first stage of attaining SA). Here, as we make Assumption 1; presented at the end of this chapter; we focus on the latter two stages of attaining SA to help us evaluate and determine the information content of the display.

Note that insufficient understanding about the situation is not solely due to lack of information in the user-interface, but it can also be the result of having an indeterministic automation or a user who is unfamiliar with system dynamics, e.g., as in category II of Pilot Induced Oscillations (PIO) [80]. In this thesis we particularly consider the case where the user is fully familiar with the deterministic dynamics of the system and our goal is to answer what information has to be presented on the display to achieve the desired task and whether it is displayed.

2.2 A framework for human-automation interaction

So far, various frameworks for analyzing human-automation interaction have been developed. Jamieson and Vicente [81] suggested a feedback loop model in which feeding back important signals to the user could help them identify and localize the source of failure in the elements of the system. Sheridan suggested a set of frameworks for systems with different levels of autonomy [82]. In [83], using a framework for supervisory control, Cummings showed that it is necessary for the user of some systems to act in collaboration with the plant, rather than exclusively acting as a supervisor.

2.2.1 Schematic framework

For this thesis, we consider a delay-free system whose evolution is modeled by an LTI model.

In Figure 2.1, we consider the user to be a function that maps the displayed information to the user input. This mapping involves different stages of information processing. In our framework, we similarly consider the user to first obtain situation awareness, then decide on the required action, and finally act on the system. As mentioned earlier, our focus is on the stage of obtaining situation awareness and our purpose is to evaluate whether or not the task is defined correctly and whether the user has access to the necessary information. We also want to determine what information is required for the user to attain SA.

We consider the user to be highly trained and experienced with the dynamics of the automation. The terms trained, experienced, and novice are frequently used in the literature to define the level of proficiency of the user, however, defining these terms formally and precisely is not an easy task. A comparison between commercial pilots, who always need to have high levels of training and experience, and licensed drivers, who are not required to have high level of proficiency, can clarify the difference between levels of training and experience for users of various systems [31]. Here by the

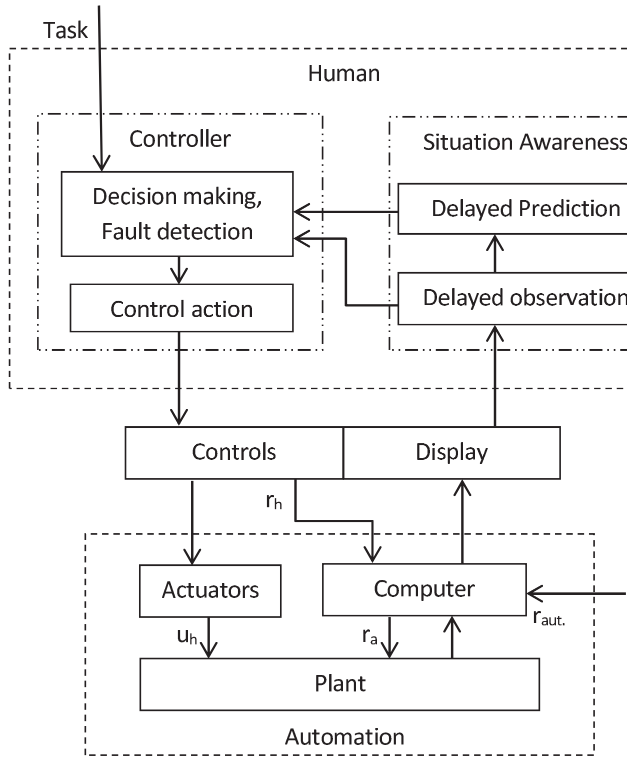


Figure 2.1: A simple schematic model of a human controlling a system. The user gets information about the automation from the display and applies control action through the controls.

terms trained and experienced, we specifically mean that the user is capable of perceiving information on the inputs, outputs, and their derivatives. We therefore ignore the information-acquisition delay of the user. The user can then make use of the perceived information to reconstruct and predict important states of the system. Recall that reconstruction (comprehension) and prediction are two stages of attaining situation awareness.

For a fully experienced and trained user with a well-developed mental model about the behavior of the system, comprehension and prediction are achieved using long term memory. This process of pattern-matching which takes place without loading the working memory is almost instantaneous.

In this thesis, we assume that the only role of the mental model is to provide the user with adequate understanding about the system's dynamics.

For the purpose of Chapter 3 we ignore the effect of any delay, however, for Chapters 4-6 we assume that the comprehension of the non-measured part of the task functional and also the prediction the task functional have to take place within the working memory which is associated with certain amount of processing delay [84]. By focusing on the processing delay and assuming the perception delay to be negligible, we build our technique on Assumption 7 for Chapters 4-6.

2.2.2 Mathematical framework

In the body of this thesis, we consider the system as a delay-free system whose evolution of the states is modeled by the LTI model

$$\dot{x}(t) = Ax(t) + Bu_h(t) + Fr_a \tag{2.1}$$

where $x(t) \in \mathbb{R}^n$ is the state vector. The inputs in (2.1) can be categorized as the known input which is the low-level human input, $u_h(t) \in \mathbb{R}^{m_b}$, controlled by the user and the input which is the time-invariant reference trajectory, $r_a \in \mathbb{R}^{m_f}$, tracked by the automation. Here, we consider the reference trajectories to be unknown, unless they are measured in the display. In (2.1), the matrices A, B, and F have compatible dimensions. Other than Chapter 3, in all other chapters we also assume that no poles of the system are on the imaginary axis.

Based on our specific application of user-interface design, we assume that each displayed measurement is either a reference trajectory or a combination of the states. Also, the output consists of two sets of measurements

$$\begin{aligned} y(t) &= Cx(t), \\ y_r(t) &= Dr_a, \end{aligned} \tag{2.2}$$

where $y \in \mathbb{R}^{p_x}$ is the set of measured combinations of states and $y_r \in \mathbb{R}^{p_r}$ is the set of measured reference trajectories in the display.

The models of the system used in the appendices are slightly different from the main model above. We will present those models in their respective chapters.

2.2.3 Task formulation

Tasks are often specified in terms of conditions that must always be met, or must be eventually met. For example, potential tasks for a remotely driven car could be *Always travel under the speed limit*, or *At some time, stop at the stop sign*. Hence we formulate the task as a function $f : \mathbb{R}^l \to \mathbb{R}^s$ with s subtasks, such that

$$\mathcal{F} = \{x \mid f(Tx) \geq 0\},\; T \in \mathbb{R}^{l \times n} \tag{2.3}$$

with task matrix T comprised of l linear combinations of the state. For *Always* $\mathcal{F}$, the state trajectory must lie in $\mathcal{F}$ for all time in order for the task to be successfully completed. For *Eventually* $\mathcal{F}$, when the state enters the set $\mathcal{F}$ at some finite time, the task has been successfully completed. For example, for a point-mass car with position x and speed v, successful completion of the task *Always travel under the speed limit* is indicated for states that remain in $\mathcal{F} = \{\underline{x} \mid v_{\text{max}} - v \geq 0\}$ for all time. Successful completion of the task *Eventually stop at the stop sign* is indicated for states that reach $\mathcal{F} = \{\underline{x} \mid x = x_{\text{stop}} \wedge v = 0\}$ at some finite time.

Denote the *task space* $\mathcal{T}$ by the row space of T, that is, $\mathcal{T} = \mathcal{R}(T^T)$.

2.3 Approaches taken in the thesis

In this thesis we perform two types of analysis to determine the correctness of the user interface content. In Chapters 3 and 4 we analyze the correctness of this information via subspace analysis and in Chapters 5 and 6 we come up with a model of the user attaining SA.

2.3.1 Subspace analysis

Degani and Heymann [79], introduced a schematic diagram to define the desired relationship between the elements of system which are the user-model, the task, and the user-interface during human-automation interaction. Using their suggested model, they could describe the interrelation between these elements of the system [79]. When user's capabilities, information from the display, and task requirements are aligned, the correct interaction between the elements of the system is possible. As an extension to their model, we suggest Figure 2.2 which presents the necessary relationship between the elements of the system.

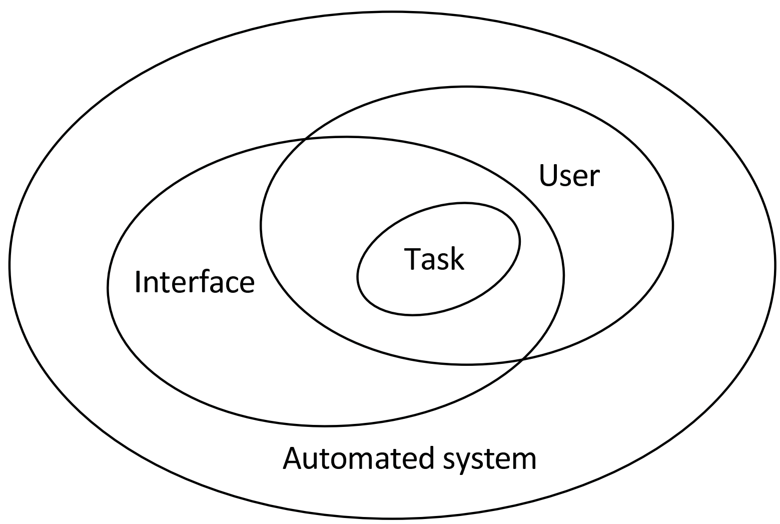


Figure 2.2: The required relationship between the elements of a system. The user have to be provided with information regarding the task in order to accomplish it.

From Figure 2.2, the human-automation interaction is not correct for all tasks unless the task requirements is entirely included in the intersection of user-model and user-interface – that is, to have a correct human-automation interaction, it is necessary for the user-interface to provide the user with information related to any feasible task.

Having developed a framework for shared control systems and having formulated the task physically and mathematically, we are ready to come-up with a mathematical criterion for evaluating the information content of the display for the system presented in Figure 2.1 to meet the requirements of a task formulated in (2.3). Hence, we should evaluate whether the user has access to the required information to attain situation awareness regarding the task. We therefore impose Assumption 5; presented at the end of this chapter.

2.3.2 The modeling technique

To have a successful HAI, both the data-driven (bottom-up) information processing and the goal-directed (top-down) processing for SA are considered to be vital [85, 86]. The emphasis of the goal-driven processing is on paying direct attention to and then processing the most important information related to the goal. Literature reviews on the relationship between attention and working memory have been provided in [87, 88]. In their discussions, [87, 88] brought evidence on how the attention acts to filter out the unnecessary information at both the early stages of perception and the late stage of processing the information in working memory. Overall, one main role of attention during the post perceptual stage of processing information is to reduce or cancel out the distractions while comprehending the target [89].

The adverse effect of irrelevant information on the ability of the user to understand and perform a desired task has been investigated extensively in the literature. Through an experiment performed over two years on a group of elementary students, [90] concluded that “ the problem-solving ability is related to the ability of reducing the memory accessibility of non-target and irrelevant information”. Similar comments were made in [91, 92], e.g., lack of the capability of selecting information relevant to the task and suppressing irrelevant information will result in poor performance of the working memory and poor comprehension. It was shown that the ability to suppress irrelevant information declines for elderly adults [93]. It was also mentioned that, for more challenging tasks, during the processing and manipulating of information in the executive working memory, the users will have higher focus on the actual goal and irrelevant task information will have less effect [88].

The above discussions suggest that it is desirable for a good user (a good problem solver) to concentrate on the target and suppress irrelevant information. Intuitively, a trained and experienced user generally will not perform unnecessary processing of information. Therefore, it is not realistic to assume that the user reconstructs all observable and predictable states to only make use of the desired functional of the states of the system. For instance, let us use an aircraft with eight lateral and longitudinal states as an example. If the desired purpose for the pilot is to keep the angle of attack bounded, intuitively s/he will not aim to reconstruct and predict all other unmeasured states of the system, e.g., the yaw rate is of no interest here. We therefore have Assumption 6; presented in Section 2.4.

Based on Assumption 6 and since we consider the comprehension and prediction of the unmeasured information to be a challenging task that can occupy the executive working memory of the user, we consider the user to behave as a functional estimator in order to process the information. We believe that this model is a reasonable model to start with, although it is simplistic as it cannot capture some common conditions such as mind wandering - that is, the decoupling of information processing from the primary task [94, 95]. The simplified model is in fact addressing perfect users who perceive all information provided in the display and can fully concentrate on what they are asked to do.

The block diagram presented in figure 2.3 contains the model of the system that we deal with and the model of the human that we suggest. The human is modeled as an observer/predictor that acts upon the information provided in the display. In this diagram, the dashed dotted blocks are the important parts of the human model which in this thesis we do not focus on (i.e. the direct effect of mental model as well as the effect of the information from the environment).

As is shown in figure 2.3, we consider the users to have access to the information about input, output, and some of their derivatives. They might also have access to the reference trajectory from the automation. This user is assumed to make delayed estimations of desired functionals of the states of the system, yet the directly measured information and its derivatives are

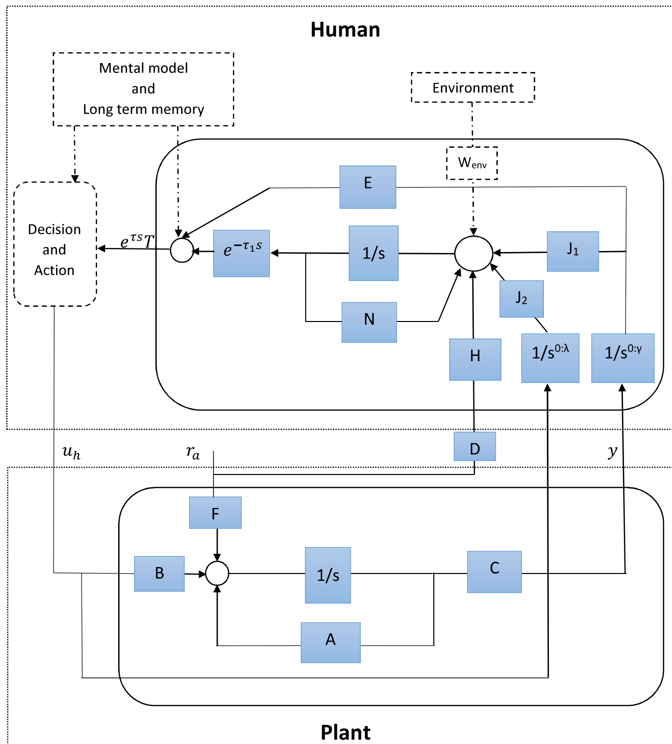


Figure 2.3: Block diagram presentation of the system and the user.

not delayed. While the input and output signals are often noisy, in this model we have not considered any noise. In addition, in real world, there might exist parametric uncertainties in the dynamics of the system which affect A, B, F, and C.

2.4 List of the assumptions

For the purpose of this thesis and based on the above discussion in the current chapter, we consider the following general assumptions:

Assumption 1. *The user can perceive all the displayed information.*

Assumption 2. *The user is provided with no additional information beyond what is on the display.*

Assumption 3. *Depending on the application, the users might have knowledge about the derivatives of their own inputs, up to λ derivatives and also have knowledge about the derivatives of the outputs, up to γ derivatives.*

Regarding the task we assume :

Assumption 4. *To assure feasibility of the task, let $\mathcal{T} \subseteq \mathcal{C}$, the controllable subspace of (2.1).*

To evaluate the correctness of the displayed information in Chapters 3 and 4, we consider:

Assumption 5. *In order to accomplish or monitor a desired task, all rows of Tx should be mathematically observable and predictable by the user.*

and for Chapters 5 and 6, we assume:

Assumption 6. *For the purpose of reconstructing the desired functional, the user does not estimate all observable states unless reconstructing all observable states is feasible and necessary for the estimation of the desired functional.*

In addition, for Chapters 4-6, we also consider:

Assumption 7. *Comprehension of the directly measured combination of the states of the system does not incorporate delay. However, prediction of all combinations of the states and also comprehension of the non-measured functionals of the states are associated with a delay, τ_1.*

Assumption 8. *Matrix A has no eigenvalues on the imaginary axis.*

Chapter 3

Novel observability /predictability subspaces to analyze user-interfaces

In this chapter, we identify basic observability-based conditions under which a user-interface provides the user with adequate information to accomplish a given task, formulated as a subset of the state-space.

We, thus, formulate the user-observable subspace in Section 3.2, and the user-predictable subspace in Section 3.3. We evaluate the mentioned subspaces for the case that the users do not know the reference trajectory, r_a, and the case that they know the pattern of changes of the reference trajectory, i.e., the reference trajectory and its derivatives are known.

3.1 Problem formulation

As has been discussed in Section 2.2.2, we model our system as in equation (2.1) such that the evolution of the states is a function of the human low level input and the automation reference trajectory. In this chapter we narrow down our investigations to the cases for which no derivative of the human input is known to the user and the pattern of changes of the output is entirely known by the user, i.e. $\lambda = 0$ and $\gamma = \infty$; this assumption will be relaxed in the next chapters.

In the formalism of Assumption 5, we say that the task should be user-observable and user-predictable, i.e., the user should be able to reconstruct the desired combination of the state of the system at the current and the next instant in time. Mathematically, in this chapter, we define a state to be user-predictable if and only if the user can reconstruct both the current value of the state and its higher derivatives, up to a pre-specified degree.

For the purpose of formulating the user-observable/predictable spaces, in chapters 3 and 4 we extensively use projection matrices. Hence, it is worthwhile to first introduce the concept of orthogonal projection onto a space through an example.

By definition, the projection matrix, P, onto the subspace $\mathcal{X} \triangleq \text{span}(X)$ is a symmetric matrix which can be computed as

$$P = X(X^T X)^{-1} X^T. \tag{3.1}$$

Lets consider the space $\mathcal{V}$ with the basis vector v such that

$$v = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

.

From (3.1), the orthogonal projection onto $\mathcal{V}$ can be obtained as

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.2}$$

and the orthogonal projection of any vector u on the space $\mathcal{V}$ is Pu.

We can also formulate the projection onto the null space of vector v by considering $X = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$. In this case the projection matrix will be

$$P = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{3.3}$$

Now, for a system of form (2.1), if we project the state vector onto the left null space of matrix B_h – that is, $\mathcal{N}({B_h}^T)$, the result will show the combinations of the states which are not affected by the input $u(t)$.

Similarly, projection onto $\mathcal{N}(F^T)$ states that the projected vector is free of the effect of r_a.

3.2 User-observable subspace

Consider system (2.1) and- the output equation in 2.2. We first assume that the automation reference trajectory is not shown to the user, i.e. in 2.2 we have $D = 0$.

The first derivative of the output equation is

$$\dot{y} = CAx + CBu_h + CFr_a, \tag{3.4}$$

with Bu_h is known to the user.

Consider an orthogonal projection matrix $P_{r,1} \in \mathbb{R}^{p\times p}$ onto $\mathcal{N}((CF)^T)$. Hence, this matrix

$$P_{r,1}\dot{y} = P_{r,1}CAx + P_{r,1}CBu_h \tag{3.5}$$

removes the unknown reference trajectory from the first derivative of the output equation. For the special case in which $m \geq p$ and $\text{rank}(CF) = p$, $P_{r,1} = 0$.

Now consider an orthogonal projection $P_{h,1} \in \mathbb{R}^{p\times p}$ onto $\mathcal{N}((P_{r,1}CB)^T)$ and an orthogonal projection $P_{r,2} \in \mathbb{R}^{p\times p}$ onto $\mathcal{N}((P_{h,1}P_{r,1}CAB)^T)$. The matrices $P_{h,1}$ and $P_{r,2}$ remove the unknown $B\dot{u}_h$ and the unknown Fr_a from the second derivative of the output equation respectively, therefore,

$$M_2\ddot{y} = M_2CA^2x + M_2CABu_h. \tag{3.6}$$

with $M_i = P_{r,i+1}\prod_{k=i}^{1} P_{h,k}P_{r,k}$ for $i = 2, \cdots, n$.

Continuing in a similar fashion with higher derivatives,

$$M_iy^{(i)} = M_iCA^ix + M_iCA^{(i-1)}Bu_h, \tag{3.7}$$

and the projection matrices are selected such that they remove the unknown values from the derivatives of output equation. Hence,

$$\begin{aligned} \prod_{k=i-1}^{1} (P_{h,k} P_{r,k}) C A^{i-2} B &= 0 \\ M_i C A^{i-1} F &= 0. \end{aligned} \tag{3.8}$$

Note that the projection matrices are chosen to remove the unknown values from the derivatives of the output equation, hence, depending on the availability of different types of inputs, the analytical expressions of these matrices may change, e.g. in Corollaries 1 and 2.

Combining the output equation, (3.5), (3.6), and (3.7), we obtain

$$\begin{bmatrix} y \\ P_{r,1}\dot{y} \\ M_2 \ddot{y} \\ \vdots \\ M_{n-1} y^{(n)} \end{bmatrix} = O_p x + \begin{bmatrix} 0 \\ P_{r,1} C B \\ M_1 C A B \\ \vdots \\ M_n C A^{n-1} B \end{bmatrix} u_h \tag{3.9}$$

with

$$O_p = \begin{bmatrix} C \\ P_{r,1} C A \\ M_2 C A^2 \\ \vdots \\ M_n C A^n \end{bmatrix}. \tag{3.10}$$

Theorem 1. *Under Assumption 3 for $\lambda = 0$ and $\gamma = \infty$, the user-observable subspace of system (2.1) with an unknown reference trajectory is*

$$\mathcal{O}_{\mathcal{H}} \triangleq \mathcal{R}(C^T) \oplus \mathcal{R}((P_{r,1} C A)^T) \oplus \sum_{i=2}^{n-1} \mathcal{R}((M_i C A^i)^T), \tag{3.11}$$

where $M_i = P_{r,i+1} \prod_{k=i}^{1} P_{h,k} P_{r,k}$ and the projection matrices can be obtained *recursively as*

$$
\begin{aligned}
P_{r,i} &= p(\mathcal{N}((\prod_{k=i-1}^{1}(P_{h,k}P_{r,k})CA^{i-1}F)^T)),\\
P_{h,i} &= p(\mathcal{N}(P_{r,i}\prod_{k=i-1}^{1}(P_{h,k}P_{r,k})CA^{i-1}B)^T)),\\
P_{r,1} &= p(\mathcal{N}((CF)^T)),\\
P_{h,1} &= p(\mathcal{N}(P_{r,1}CB)^T)),
\end{aligned}
\tag{3.12}
$$

where $p(\mathcal{M})$ means the projection onto the space $\mathcal{M}$.

Proof. Given that the only unknown values in (3.9) is the state vector x, the combinations of the state which span the row space of of O_p can be reconstructed by the user. Hence,

$$
\begin{aligned}
\mathcal{R}({O_p}^T) = \ & \mathcal{R}(C^T) \oplus \mathcal{R}((P_{r,1}CA)^T)\oplus \\
& \mathcal{R}((M_2CA^2)^T) \oplus \cdots \oplus \mathcal{R}((M_nCA^n)^T).
\end{aligned}
\tag{3.13}
$$

Equation (3.13) and equivalently the right hand side of (3.11) represent a subspace that can be reconstructed by the user for $\lambda = 0$ and $\gamma = \infty$ – that is, $O_p \triangleq \mathcal{O}_\mathcal{H}$. □

Note that a *user-observable* system is a system for which $\mathcal{O}_\mathcal{H} = \mathbb{R}^n$.

Corollary 1. *Under Assumption 3, the user-observable subspace of system (2.1) with a known reference trajectory is*

$$
\mathcal{O}_\mathcal{H} \triangleq \mathcal{R}(C^T) \oplus \sum_{i=1}^{n-1}\mathcal{R}((\prod_{k=i}^{1}P_{h,k}CA^i)^T),
\tag{3.14}
$$

where

$$
\begin{aligned}
P_{h,i} &\triangleq p(\mathcal{N}((\textstyle\prod_{k=i-1}^{1}P_{h,k}CA^{i-1}B)^T)),\\
P_{h,1} &\triangleq p(\mathcal{N}((CB)^T)).
\end{aligned}
\tag{3.15}
$$

Proof. Similar to the proof of Theorem 1, we now consider that r_a and its derivatives are available, therefore, there is no need for projection matrices to remove the reference trajectory from the derivatives of states equations. $\square$

Corollary 2. *Under Assumption 3 for $\lambda = 0$ and $\gamma = \infty$, the user-observable subspace of*

- *the human-driven system is*

$$\mathcal{O}_{\mathcal{H}} \triangleq \mathcal{R}(C^T) \oplus \sum_{i=1}^{n-1} \mathcal{R}((\prod_{k=i}^{1} P_{h,k} C A^i)^T), \tag{3.16}$$

  where

$$\begin{array}{rcl} P_{h,i} & \triangleq & p(N((\prod_{k=i-1}^{1} P_{h,k} C A^{i-1} B)^T)), \\ P_{h,1} & \triangleq & p(\mathcal{N}((CB)^T)). \end{array} \tag{3.17}$$

- *the systems under reference tracking is equivalent to the observable subspace of those systems.*

$$\mathcal{O}_{\mathcal{H}} \triangleq \mathcal{R}(C^T) \oplus \sum_{i=1}^{n-1} \mathcal{R}((\prod_{k=i}^{1} P_{r,i} C A^i)^T), \tag{3.18}$$

  where

$$\begin{array}{rcl} P_{r,i} & \triangleq & p(N((\prod_{k=i-1}^{1} P_{r,k} C A^{i-1} F)^T)), \\ P_{r,1} & \triangleq & p(\mathcal{N}((CF)^T)). \end{array} \tag{3.19}$$

Proof. Consider that for a human-driven system $F = 0$. In addition, we assume that the reference trajectory is unknown in the case of monitoring the states of an automated system, but, for reference tracking it is assumed to be known. Hence, the user-observable subspace of different paradigms of human-automation interaction can be easily obtained from (3.11) and (3.14). $\square$

3.3 User-predictable subspace

To obtain the user-predictable space, we define a transformation matrix

$$S \triangleq \left[\begin{array}{c} {E_{\mathcal{O}_\mathcal{H}}}^T \\ \hline {E^{\perp}_{\mathcal{O}_\mathcal{H}}}^T \end{array}\right], \tag{3.20}$$

with $E_{\mathcal{O}_\mathcal{H}}$ the basis of the user-observable subspace of (2.1) and $E^{\perp}_{\mathcal{O}_\mathcal{H}}$ any orthogonal complement of $E_{\mathcal{O}_\mathcal{H}}$. The space $E_{\mathcal{O}_\mathcal{H}}$ and therefore $E^{\perp}_{\mathcal{O}_\mathcal{H}}$ can be easily obtained from the results of Theorem 1. Using $\bar{x} \triangleq Sx$ the state vector breaks into user-observable states, $\bar{x}_{\mathcal{O}_\mathcal{H}}$, and user-unobservable states, $\bar{x}_{\mathcal{U}\mathcal{O}_\mathcal{H}}$. The transformed system is

$$\left[\begin{array}{c} \dot{\bar{x}}_{\mathcal{O}_\mathcal{H}} \\ \hline \dot{\bar{x}}_{\mathcal{U}\mathcal{O}_\mathcal{H}} \end{array}\right] = \bar{A} \left[\begin{array}{c} \bar{x}_{\mathcal{O}_\mathcal{H}} \\ \hline \bar{x}_{\mathcal{U}\mathcal{O}_\mathcal{H}} \end{array}\right] + \bar{B}u_h + \bar{F}r_a \tag{3.21}$$

with $\bar{A} = SAS^{-1}$, $\bar{B} = SB$, $\bar{F} = SF$, $\bar{C} = CS^{-1}$, and $\bar{D} = DS^{-1}$.

$$\begin{aligned} \bar{A} &= \left[\begin{array}{c|c} \bar{A}g_{\mathcal{O}_\mathcal{H}} & \bar{A}g_{12} \\ \hline \bar{A}g_{21} & \bar{A}g_{\mathcal{U}\mathcal{O}_\mathcal{H}} \end{array}\right], \\ \bar{F} &= \left[\begin{array}{c} \bar{B}g_{\mathcal{O}_\mathcal{H}} \\ \hline \bar{B}g_{U\mathcal{O}_\mathcal{H}} \end{array}\right], \bar{B} = \left[\begin{array}{c} \bar{B}_{H\mathcal{O}_\mathcal{H}} \\ \hline \bar{B}_{HU\mathcal{O}_\mathcal{H}} \end{array}\right]. \end{aligned} \tag{3.22}$$

From (3.21), the derivatives of the user-observable states can be calculated

$$\begin{aligned} {E_{\mathcal{O}_\mathcal{H}}}^T \dot{x} &= \left[\begin{array}{c|c} \bar{A}g_{\mathcal{O}_\mathcal{H}} & \bar{A}g_{12} \end{array}\right] \left[\begin{array}{c} \bar{x}_{\mathcal{O}_\mathcal{H}} \\ \hline \bar{x}_{\mathcal{U}\mathcal{O}_\mathcal{H}} \end{array}\right] \\ &\quad + \bar{B}_{H\mathcal{O}_\mathcal{H}} u_h + {E_{\mathcal{O}_\mathcal{H}}}^T F r_a \end{aligned} \tag{3.23}$$

with $\bar{x}_{\mathcal{U}\mathcal{O}_\mathcal{H}}$ and the unknown reference trajectory preventing the user from reconstructing the derivative of user-observable states.

Consider two projection matrices $P_{x,1}$ and $P_{r,1}$ with the property to remove the unknown values from (3.23) – that is, $P_{x,1}$ is orthogonal projection onto $\mathcal{N}(\bar{A}g_{1,2})$ and $P_{r,1}$ is an orthogonal projection on to $\mathcal{N}((P_{x,1}E_{\mathcal{O}_\mathcal{H}}{}^T F)^T)$. Hence, from

$$P_{r,1}P_{x,1}E_{\mathcal{O}_\mathcal{H}}{}^T\dot{x} = P_{r,1}P_{x,1}\bar{A}g_{\mathcal{O}_\mathcal{H}}\bar{x}_{\mathcal{O}_\mathcal{H}} + P_{r,1}P_{x,1}\bar{B}_{H\mathcal{O}_\mathcal{H}}u_h, \tag{3.24}$$

the first derivative of the states which spans $\mathcal{R}((P_{r,1}P_{x,1}E_{\mathcal{O}_\mathcal{H}}{}^T)^T)$ can be reconstructed.

Continue in a similar fashion and take higher derivatives from the state equation. At each stage introduce three projection matrices to remove the unknown values, r_a, $\bar{x}_{\mathcal{U}\mathcal{O}_\mathcal{H}}$, and $\bar{B}_{H\mathcal{O}_\mathcal{H}}\dot{u}_h$. Therefore, with

$$N_i = \prod_{k=i}^{2}(P_{h,k-1}P_{r,k}P_{x,k})P_{r,1}P_{x,1},$$

the i^{th} derivative of state can be obtained from

$$N_iE_{\mathcal{O}_\mathcal{H}}{}^T x^{(i)} = N_i\bar{A}g^{i}_{\mathcal{O}_\mathcal{H}}\bar{x}_{\mathcal{O}_\mathcal{H}} + N_i\bar{A}g^{i-1}_{\mathcal{O}_\mathcal{H}}\bar{B}_{H\mathcal{O}_\mathcal{H}}u_h. \tag{3.25}$$

Theorem 2. *For the LTI system (2.1) with an unknown reference trajectory and under Assumption 3 for $\lambda = 0$ and $\gamma = \infty$, the user-predictable subspace is*

$$\mathcal{P}_\mathcal{H} \triangleq \bigcap_{i=1}^{n_p-1}\mathcal{R}((N_iE_{\mathcal{O}_\mathcal{H}}{}^T)^T) \tag{3.26}$$

where n_p shows the degree of derivatives of the states which need to be available for their predictability, $N_i = \prod_{k=i}^{2}(P_{h,k-1}P_{r,k}P_{x,k})P_{r,1}P_{x,1}$, and

$$\begin{array}{rcl} P_{x,i} & = & p(\mathcal{N}((N_{i-1}\bar{A}g^{i-1}_{\mathcal{O}_\mathcal{H}}\bar{A}g_{1,2})^T)), \\ P_{r,i} & = & p(\mathcal{N}((P_{x,i}N_{i-1}\bar{A}g^{i-1}_{\mathcal{O}_\mathcal{H}}E_{\mathcal{O}_\mathcal{H}}{}^T F)^T)), \\ P_{h,i} & = & p(\mathcal{N}((P_{r,i}P_{x,i}N_{i-1}\bar{A}g^{i-1}_{\mathcal{O}_\mathcal{H}}E^T_{\mathcal{O}_\mathcal{H}}B)^T)) \end{array} \tag{3.27}$$

and

$$\begin{aligned} P_{x,1} &= p(\mathcal{N}(\bar{A}g_{1,2}^T)), \\ P_{r,1} &= p(\mathcal{N}((P_{x,1}E_{\mathcal{O_H}}{}^T F)^T)), \\ N_1 &= P_{r,1}P_{x,1}, \end{aligned} \tag{3.28}$$

such that k is the highest derivative which should be observable by the user and $p(M)$ means the projection onto the space $\mathcal{M}$.

Proof. From (3.24), the user can reconstruct the derivative of those combinations of states which span $\mathcal{R}(E_{\mathcal{O_H}}P_{x,1}^T P_{r,1}^T)$. Similarly, from (3.25), the user can reconstruct the i^{th} derivative of those combinations of states which span $\mathcal{R}(E_{\mathcal{O_H}}N_i{}^T)$. Therefore, in order to reconstruct higher derivatives of state vector (up to the k^{th} derivative), the user predictable subspace can be obtained as equation (3.26) and Theorem 2 is proved. $\square$

Note that a *user-predictable* system is the one that satisfies $\mathcal{P_H} = \mathbb{R}^n$.

Corollary 3. *Under Assumptions 3 for $\lambda = 0$ and $\gamma = \infty$, the user-predictable subspace of system (2.1) with a known reference trajectory is*

$$\mathcal{P_H} \triangleq \bigcap_{i=1}^{n_p-1} \mathcal{R}((N_i E_{\mathcal{O_H}}{}^T)^T) \tag{3.29}$$

where $N_i = \prod_{k=i}^{2}(P_{h,k-1}P_{x,k})P_{x,1}$ *and*

$$\begin{aligned} P_{x,i} &= p(\mathcal{N}((N_{i-1}\bar{A}g_{\mathcal{O_H}}^{i-1}\bar{A}g_{1,2})^T)), \\ P_{h,i} &= p(\mathcal{N}((P_{x,i}N_{i-1}\bar{A}g_{\mathcal{O_H}}^{i-1}E_{\mathcal{O_H}}^T B)^T)) \end{aligned} \tag{3.30}$$

and

$$\begin{aligned} P_{x,1} &= p(\mathcal{N}(\bar{A}g_{1,2}^T)), \\ N_1 &= P_{x,1}, \end{aligned} \tag{3.31}$$

such that k the highest derivative which should be observable by the user.

Proof. The proof is trivial. Consider that when the reference trajectory and its derivatives are all available, there is no need for an orthogonal projection matrix to remove them from the derivatives of the output equation. $\square$

3.4 Examples

3.4.1 Nagoya A300 Accident, 1994

For the Airbus accident of Section 1.1, we model the aircraft prior to stall as an LTI system under shared control, and aim to determine whether the pilots had access to the required information about the aircraft to assure predictability of the states relevant to the desired task.

Consider the linearized longitudinal dynamics of an aircraft in trimmed level flight, with state $x = [V, \alpha, q, \theta]$ consisting of total velocity V, angle of attack α, pitch rate q and pitch angle θ and state matrix

$$A = \begin{bmatrix} -0.0414 & 10.9259 & 0 & -32.8000 \\ -0.0013 & -0.5017 & 1.0000 & 0 \\ 0 & -0.4184 & 0 & 0 \\ 0 & 0 & 1.0000 & 0 \end{bmatrix}, \tag{3.32}$$

with the numerical values for stability derivatives taken from data for the Boeing 747 [96]. The equations of the motion of the aircraft from which matrix A is achieved are provided in [97].

In order to achieve and keep a desired altitude, we consider that the automation is following a glide slope which approaches zero, i.e. $C_d = [1\ 0\ -1\ 0]$. Hence, both the pilot and the autopilot were trying to control the flight path angle, one in order to descend and the other one in order to ascend the aircraft. The accident investigation report [5] states that the pilot and the co-pilot expected their control input to override the autopilot control action, which in fact was not a valid expectation, therefore, we consider the reference trajectory to be unknown to the user. In Airbus A300, the pilots control the flight path angle using the elevators,

$$B = \begin{bmatrix} 0 & -0.0305 & -0.4039 & 0 \end{bmatrix}^T \tag{3.33}$$

and the autopilots control it through the angle of the horizontal tail (THS), thus, the automation affects the states through

$$B_a = \begin{bmatrix} 0 & -0.0636 & -0.8656 & 0 \end{bmatrix}^T, \tag{3.34}$$

The input δ_e and δ_{ih} represent elevator deflection and THS deflection respectively. Having B_a from (3.34), it is straight forward to obtain F.

Velocity, pitch angle, and flight path angle are all available in standard A300-600 pilot displays [98]. Hence, for the augmented system

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \end{bmatrix}. \tag{3.35}$$

We define the desired task of the pilot to be descending on the desired glideslope, $\gamma = \gamma_{des}$. Since $\gamma = \alpha - \theta$, by (2.3), we obtain

$$\mathcal{T} = \text{span}\left(\begin{bmatrix} 0 & 1 & 0 & -1 \end{bmatrix}^T\right). \tag{3.36}$$

By applying Theorem 1, we obtain the user-observable subspace $\mathcal{O}_{\mathcal{H}} = \text{span}(e_1, \cdots, e_4)$ which shows that by measuring the mentioned states of the system, the pilot would be able to reconstruct the unknown pitch rate.

We assume that the first two derivatives of user-observable states should be observable in order to achieve a reliable prediction of their values. From (3.26), the user predictable subspace is

$$\mathcal{P}_{\mathcal{H}} = \text{span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ -0.8251 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -0.0735 \\ -0.9644 \end{bmatrix}\right). \tag{3.37}$$

Since $\mathcal{T} \nsubseteq \mathcal{P}_{\mathcal{H}}$, our method demonstrates that the modeled user-interface does not provide the necessary information for the user to predict the system's behavior.

If we assume that the automation's desired trajectory is known to the user, still the flight path angle and, as a result, the angle of attack will not be user-predictable. Hence, using our method could be helpful in the early stages of the design of flight management systems by showing the designers that in some specific modes of the flight, flying under shared control, e.g., shared control on the flight-path angle, can be hazardous even for pilots who are aware of the autopilot's goal.

3.4.2 A remotely controlled fleet of UAVs

Several researchers have developed methods for stabilization and control of the formation of a group of UAVs for both continuous-time and discrete-event systems [99–101]. We consider a fleet of two point mass vehicles, flying in a leader-follower formation and following a real-time reference trajectory defined by a remote operator. We model each vehicle as a double integrator,

$$A_i = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B_i = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{3.38}$$

In the leader-follower formation, each vehicle only has access to the information about the position of its leader and the main leader is following a reference trajectory.

We model the formation such that the position of each vehicle is the reference trajectory for its direct follower and the overall trajectory of the formation is the one which is defined by the user to the leader of the formation. Hence,

$$\begin{aligned} A &= \left[\begin{array}{c|c} A_i & 0 \\ \hline B & A_i \end{array}\right], \\ B &= \left[\begin{array}{c|c} {B_i}^T & 0 \end{array}\right]^T. \end{aligned} \tag{3.39}$$

Under Assumption 3 and based on the availability of human reference trajectory we can evaluate the user-observable and the user-predictable subspaces for different measurements from Corollary 2 and 3. Table 3.1 shows how the various measurements can affect the mentioned subspaces.

As opposed to what one might expect, Table 3.1 demonstrates that providing information about the position of the leader does not let the user

Measurement	C	$\mathcal{O}$, $\mathcal{O}_{\mathcal{H}}$, and $\mathcal{P}_{\mathcal{H}}$
Leader position	${e_1}^T$	$\text{span}(e_1, e_2)$
Follower position	${e_3}^T$	$\mathbb{R}^4$

Table 3.1: The observable and predictable subspaces for different measurements of a leader-follower formation

comprehend and predict the position of the follower. Essentially, by measuring the states of the leader the user may only be capable of accomplishing tasks related to the states of the leader. The reason lies behind the fact that in the leader-follower formation the states of the leader are not affected by its follower – that is, changes in the states of the leader are only made by the operator's command.

The states of the follower are on the other hand affected by the states of its leader. From Table 3.1, in order to accomplish any possible set of tasks in a leader-follower formation, i.e., in order to satisfy the inclusion $\mathcal{T} \subseteq \mathcal{P}_{\mathcal{H}}$ for all possible sets of tasks, the position of the follower has to be measured in the user interface.

3.5 Discussion and an alternative presentation of the system

In this chapter we modeled the evolution of the system as a function of the low-level human input and the automation's reference trajectory. An alternative presentation of the system with the relevant analysis is provided in our conference paper [102] in which we consider the states of the system to be a function of low-level inputs from both the user and the automation. In the alternative presentation of the system, we consider the system to be a non delayed continuous time LTI system whose evolution is modeled as

$$\begin{aligned} \dot{x} &= Ax + \begin{bmatrix} B_h & B_a & B_{both} \end{bmatrix} \begin{bmatrix} u_h \\ u_a \\ u_{both} \end{bmatrix} \\ y &= Cx \end{aligned} \tag{3.40}$$

with $A \in \mathbb{R}^{n \times n}$, $B_h \in \mathbb{R}^{n \times m_h}$, $B_a \in \mathbb{R}^{n \times m_a}$, $B_{both} \in \mathbb{R}^{n \times m_b}$, $C \in \mathbb{R}^{p \times n}$, state vector $x(t) \in \mathbb{R}^n$, output $y(t) \in \mathbb{R}^p$, human input $u_h(t) \in \mathbb{R}^{m_h}$, automation input $u_a(t) \in \mathbb{R}^{m_a}$, and merged input $u_{both}(t) \in \mathbb{R}^{m_b}$. We assume u_h, u_a and u_{both} to be exhaustive and mutually exclusive inputs to the system, which model the effect of different actuators. In contrast to the standard LTI model [103], the input is categorized in one of three ways: as an input u_h controlled solely by the user, an input u_a controlled solely by the automation, or as an input u_{both} which controlled by both the user and automation. Hence, B_h models how the human input affects the system, B_a models how the automation input affects the system, and B_{both} models how the merged control input, in which commands from the human and the automation are combined in some fashion (not specified here), affects the system. Since the automation input and intent are unknown to the user, we considered u_a, u_{both}, and their derivatives, to be unknown.

Detailed discussion on how we obtained the user-observable and the user-predictable subspaces of such system is provided in [102].

3.6 Summary

This chapter presented necessary conditions for evaluating the information content of a user-interface for an LTI system under shared control. Two subspaces, the user-observable subspace $\mathcal{O}_{\mathcal{H}}$ and the user-predictable subspace $\mathcal{P}_{\mathcal{H}}$, were formulated. The user-predictable subspace is compared to the task subspace. If the task subspace does not lie in the user-predictable subspace, then the user-interface is not correct, meaning that it is not possible for the human to accomplish the desired task with the information provided on the given user interface.

The results of this chapter are acceptable for the systems in which a small amount of delay in comprehension and prediction of the task vector does not affect the safety of the system. However, for safety-critical systems which have to follow a precise trajectory of the functionals of the states, these results are not precise enough. In addition, here we have assumed that a subsequent value of a functional of the states is known if the current functional and its derivative are known. By relaxing these assumptions, we will be able to modify our tool so that it can be used for the analyses of the displayed information in safety-critical systems. In the next chapter, we modify the above observability and predictability subspaces by considering the perception delay as well as by looking at longer term predictions rather than the instantaneous prediction.

Chapter 4

Novel observability /predictability subspaces considering the delay and long term prediction

In this chapter, for a system of the form (2.1), we modify the results of Chapter 3 by taking into account the processing delay of the estimation and the prediction which we simply ignored in the previous chapter. In addition, in Chapter 3 we assumed that a subsequent value of a functional of the states is known if the current functional and its derivatives are known. Here, we relax this assumption and consider the actual evolution of the states of the system which depends on the current functional, the inputs, and input derivatives. By relaxing these assumptions, we will be able to modify our tool so that it can be used for the analyses of the displayed information in safety-critical systems.

We determine formulas for the delay-incorporating user-observable subspace, and the delay-incorporating user-predictable subspace of shared control systems in Section 4.2 and 4.3 respectively. An example on a remotely driven car is provided in Section 4.4.

4.1 Problem formulation

Under Assumption 5, the user-interface of a safety-critical system under shared control must provide the user with information that results in a delay-incorporating user-observable and a delay-incorporating user-predictable subspaces which we define them as below.

Definition 2. *The delay-incorporating user-observable subspace, $\mathcal{O}^*_{\mathcal{H}}$, is a space which is spanned by the combination of current states, $x(t)$, which are known to the user at current time, t.*

Definition 3. *The delay-incorporating user-predictable subspace, $\mathcal{P}^*_{\mathcal{H}}$, is a space which is spanned by the combination of upcoming states, $x(t+\tau)$, which are known to the user at current time, t.*

To make this more clear, consider a user who aims to reconstruct a combination of the states $x(t)$ given $u_h(t)$ and $y(t)$. Although this reconstruction might be feasible for the user, due to the processing delay, the desired combination of $x(t)$ will become available to the user at time $t+\tau_1$. For a safety critical system, this late understanding about the states of the system can be unacceptable.

Under Definitions 2 and 3, for system (2.1), we formulate the delay-incorporating user-observable subspace, and the delay-incorporating user-predictable subspace.

From Assumption (7), we can write,

$$\mathcal{O}^*_{\mathcal{H}} = \mathcal{O}_{\mathrm{H,y}} \oplus \mathcal{O}_{\mathrm{H},\tau_1} \tag{4.1}$$

where $\mathcal{O}^*_{\mathcal{H}}$ is the delay-incorporating user-observable space, $\mathcal{O}_{\mathrm{H,y}} \triangleq \mathcal{R}(C^T)$ is obtained from the directly measured combinations of the states, $y(t) = Cx(t)$, and $\mathcal{O}_{\mathrm{H},\tau_1}$ is the delayed observable space which is the space spanned by the non-measured functional of states which can be reconstructed given $\tau_1 \geq 0$ delay.

4.2 Delay-incorporating user-observable subspace

In this section, we determine the combination of the states at time t, i.e. $x(t)$, which can be reconstructed by time t. Mathematically, this means to obtain what combination of $x(t)$, can be reconstructed given $y(t-\tau_1)$, $u_h(t-\tau_1)$, and some of their derivatives.

Consider the output equation and its i^{th} derivative for $i \in \{1, \cdots, \gamma\}$ at time $t-\tau_1$,

$$\begin{aligned} y(t-\tau_1) &= Cx(t-\tau_1), \\ y^{(i)}(t-\tau_1) &= CA^{(i)}x(t-\tau_1) + CA^{i-1}Bu_h(t-\tau_1) + \\ & \cdots + CB{u_h}^{(i-1)}(t-\tau_1) + CA^{i-1}Fr_a. \end{aligned} \tag{4.2}$$

putting all the derivatives of the output equation together in a matrix form we obtain

$$Y_{0:\gamma}(t-\tau_1) = Ox(t-\tau_1) + H_x U_{0:\gamma}(t-\tau_1) + H_r r_a, \tag{4.3}$$

with

$$\begin{aligned} Y_{0:\gamma}(t-\tau_1) &\triangleq \begin{bmatrix} y(t-\tau_1) \\ \dot{y}(t-\tau_1) \\ \vdots \\ y^{(\gamma)}(t-\tau_1) \end{bmatrix}, \\ U_{0:\gamma}(t-\tau_1) &\triangleq \begin{bmatrix} u_h(t-\tau_1) \\ \dot{u}_h(t-\tau_1) \\ \vdots \\ {u_h}^{(\gamma)}(t-\tau_1) \end{bmatrix}. \end{aligned} \tag{4.4}$$

In (4.3), O is the observability matrix and H_x is the Toeplitz matrix obtained from (4.2), in addition,

$$H_r \triangleq \begin{bmatrix} 0 \\ CAF \\ \vdots \\ CA^{\gamma-1}F \end{bmatrix}. \tag{4.5}$$

In equation (4.3), the delayed-states of the system – that is, $x(t-\tau_1)$, are formulated as a function of the input, output, their derivatives up to the

γth derivative, and the reference trajectory at time $t-\tau_1$.

From Assumption 3, only λ derivatives of the input is known to the user. Hence, it is not possible to directly obtain $x(t-\tau_1)$ from (4.3). We, therefore, define $P_0 = I$ and

$$P_i = \prod_{j=0}^{\min(i-1,\lambda+1)} P_{i,j} P_{i,r}. \tag{4.6}$$

If we select $P_{i,r}$ and $P_{i,j}$ as follows, pre-multiplication of the i^{th} derivative of the output equation by $\prod_{k=0}^{i} P_i$ will remove the unknown values from the i^{th} derivative of output.

- The matrix $P_{i,r}$ is the projection matrix onto the left null-space of $D^c(\prod_{k=0}^{i-1} P_k C A^{i-1} F)$, where $D^c \in \{0, I\}$ is the complement of D – that is, $D^c + D = I$ all the time.
- For $j \leq \lambda$, $P_{i,j} = I$ and for $j = \lambda+1$ where $i > j$, $P_{i,j}$ is the projection onto the left null-space of $(P_{i,r} \prod_{k=0}^{i-1} P_k C A^{i-j-1} B)$.

Now, consider a matrix

$$M \triangleq \begin{bmatrix} P_0 & 0 & 0 & 0 \\ 0 & P_1 P_0 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \prod_{i=0}^{\gamma} P_i \end{bmatrix}, \tag{4.7}$$

and pre-multiply it in (4.3) to eliminates the unknown values of the input derivatives and the reference trajectory. Hence,

$$MY(t-\tau_1) = MOx(t-\tau_1) + MH_x U_{0:\lambda}(t-\tau_1) + MH_r r_a \tag{4.8}$$

in which the only unknown value is $x(t-\tau_1)$.

Theorem 3. *In a system of form (2.1) and under the Assumption 3, the delay-incorporating user-observable subspace is of the form*

$$\mathcal{O}^*_{\mathcal{H}} \triangleq \mathcal{R}(C^T) \oplus \mathcal{R}((e^{-A\tau_1})^T (MO)^T P_{\tau_1}^T), \tag{4.9}$$

where M is from (4.7) and P_{τ_1} is from (4.15).

Proof. Since in general the delay is small, we consider τ_1^2 to be negligible, hence,

$$u_h(t) = u_h(t-\tau_1) + \tau_1 \dot{u}_h(t-\tau_1). \tag{4.10}$$

Note that to make the results more precise, it is straight forward to model the current state as a larger series of previous states and modify the rest of the results as per need.

Under (4.10), the states of a continuous time system (2.1) evolve as

$$x(t) = e^{A\tau_1} x(t-\tau_1) + c, \tag{4.11}$$

where

$$c = \int_{t-\tau_1}^{t} e^{A(t-T)} [B \quad F] \begin{bmatrix} u_h(T) \\ r_a \end{bmatrix} dT. \tag{4.12}$$

By introducing the variables

$$\begin{aligned} \theta_0 &\triangleq (e^{A\tau_1} - I)A^{-1}B, \\ \theta_1 &\triangleq ((e^{A\tau_1} - I)A^{-1} - \tau_1 I)A^{-1}B, \\ \theta_2 &\triangleq (e^{A\tau_1} - I)A^{-1}F. \end{aligned} \tag{4.13}$$

From (4.10) - (4.13), we can write

$$x(t-\tau_1) = e^{-A\tau_1} x(t) - e^{-A\tau_1} \prod_{k=0}^{1} \theta_k {u_h}^{(k)}(t-\tau_1) - e^{-A\tau_1}\theta_2 r_a. \tag{4.14}$$

We can combine (4.8) and (4.14) to formulate $x(t)$ as a function of the input, output and their derivatives at time $t-\tau_1$. As, new unknown $u_h(t-$

τ_1), $\dot{u}_h(t-\tau_1)$, and r_a may arise, we introduce

$$P_{\tau_1} \triangleq \prod_{k=0}^{1} P_{\tau_1,k} P_{\tau_1,r} \tag{4.15}$$

where $P_{\tau_1,0}$, $P_{\tau_1,1}$, and $P_{\tau_1,r}$ are defined as follows.

- The matrix $P_{\tau_1,r}$ is a projection onto the left null-space of $D^c(MOe^{-A\tau_1}\theta_1)$.
- For $j \leq \lambda$, $P_{\tau_1,j}$ is an identity matrix and it is a projection onto the left null-space of $(\prod_{k=0}^{j-1} P_{\tau_1,k} P_{\tau_1,r} MOe^{-A\tau_1}\theta_k)$ otherwise.

From (4.14) and (4.15) we can rewrite (4.8) as

$$P_{\tau_1} MY_{0:\gamma}(t-\tau_1) = P_{\tau_1} MOe^{-A\tau_1}x(t) - C_{known} \tag{4.16}$$

where

$$\begin{aligned} C_{known} \quad = \quad & P_{\tau_1} e^{-A\tau_1}\theta_0 u_h(t-\tau_1) + P_{\tau_1} e^{-A\tau_1}\theta_1 \dot{u}_h(t-\tau_1) + \\ & P_{\tau_1} e^{-A\tau_1}\theta_2 r_a - P_{\tau_1} MH_x U_{0:\lambda}(t) - P_{\tau_1} MH_r r_a. \end{aligned} \tag{4.17}$$

Hence, the combination of $x(t)$ which spans $\mathcal{R}((P_{\tau_1} MOe^{-A\tau_1})^T)$ can be reconstructed from $Y_{0:\gamma}(t-\tau_1)$ and $U_{0:\lambda}(t-\tau_1)$. Therefore, from (4.1), Theorem (3) is proved. $\square$

Procedure. The following steps are required to calculate the delay-incorporating user-observable subspace, $\mathcal{O}^*_{\mathcal{H}}$:

- Determine matrices O, H_x, and H_r and calculate the value of θ_0-θ_2 from (4.13).
- Obtain P_i from (4.6) and P_{τ_1} from (4.15).
- Determine M from (4.7).
- Determine the delay-incorporating user-observable subspace from (4.9).

$\square$

Corollary 4. *A delay-incorporating user-observable space is also user-observable.*

Proof. By definition, the user-observable space is the space which is spanned by the combination of the current states which are known to the user at the current time, for $\tau_1 = 0$.

From (4.8), the user-observable subspace can be formulated as

$$\mathcal{O}_{\mathcal{H}} \triangleq \mathcal{R}((MO)^T). \tag{4.18}$$

We also have the equation of $\mathcal{O}^*_{\mathcal{H}}$ from (4.9).

Under Assumption (8), the matrix A is of full rank, hence

$$\mathcal{R}((e^{-A\tau_1})^T(MO)^T P^T_{\tau_1}) = \mathcal{R}((MO)^T P^T_{\tau_1}).$$

In addition, for random matrices N and Q of compatible dimensions, we have $\mathcal{R}(NQ) \subseteq \mathcal{R}(N)$. Hence,

$$\begin{aligned} \mathcal{O}^*_{\mathcal{H}} &\triangleq \mathcal{R}((e^{-A\tau_1})^T(MO)^T P^T_{\tau_1}) \\ &= \mathcal{R}((MO)^T P^T_{\tau_1}) \\ &\subseteq \mathcal{R}((MO)^T), \end{aligned} \tag{4.19}$$

which proves that $\mathcal{O}^*_{\mathcal{H}} \subseteq \mathcal{O}_{\mathcal{H}}$. $\square$

4.3 Delay-incorporating user-predictable subspace

From Definition (3), the delay-incorporating user-predictable space is the space which can be spanned at time t based on the information available on $x(t+\tau)$. As in Section 4.2, we can write the upcoming states as a function of $Y_{0:\gamma}(t-\tau_1)$ and $U_{0:\lambda}(t-\tau_1)$.

Theorem 4. *In a system of form (2.1) and under the Assumption 3, the delay-incorporating user-predictable subspace is of the form*

$$\mathcal{P}^*_{\mathcal{H}} \triangleq \mathcal{R}((e^{-A(\tau+\tau_1)})^T(MO)^T P^T_{\tau_1} P^T_{\tau}), \tag{4.20}$$

where P_τ is from (4.24), P_{τ_1} is from (4.15), and M is from (4.7).

Proof. From (4.16) and

$$x(t) = e^{-A\tau}x(t+\tau) - e^{-A\tau}\sum_{i=0}^{1}\delta_i u_h(t)^{(i)} - e^{-A\tau}\delta_2 r_a \tag{4.21}$$

where τ is the required prediction horizon and

$$\begin{aligned} \delta_0 &\triangleq (e^{A\tau} - I)A^{-1}B, \\ \delta_1 &\triangleq ((e^{A\tau} - I)A^{-1} - \tau I)A^{-1}B, \\ \delta_2 &\triangleq (e^{A\tau} - I)A^{-1}F, \end{aligned} \tag{4.22}$$

we can write

$$\begin{aligned} P_{\tau_1}MY_{0:\gamma}(t-\tau_1) = \; & P_{\tau_1}MOe^{-A(\tau_1+\tau)}x(t+\tau) - P_{\tau_1}MOe^{-A\tau}\delta_1 u_h(t-\tau_1) - \\ & P_{\tau_1}MOe^{-A\tau}(\delta_2+\tau_1\delta_1)\dot{u}_h(t-\tau_1) - P_{\tau_1}MOe^{-A\tau}\delta_3 r_a - C_{known} \end{aligned} \tag{4.23}$$

where C_{known} is from (4.17).

By pre-multiplying (4.23) by

$$P_\tau \triangleq \prod_{k=0}^{1} P_{\tau,j}P_{\tau,r} \tag{4.24}$$

we can remove all unknown values from it. In (4.24),

- The matrix $P_{\tau,r}$ is a projection onto the left null-space of $D^c(P_{\tau_1}MOe^{-A(\tau_1+\tau)}\delta_2)$.
- For $j \leq \lambda$, $P_{\tau,j}$ is an identity matrix and it is a projection onto the left null-space of $(\prod_{k=0}^{j-1} P_{\tau,j}P_{\tau,r}P_{\tau_1}MOe^{-A(\tau_1+\tau)}\delta_k)$ otherwise.

Hence, the functional of the upcoming states of the system, $x(t+\tau)$, which span the row space of $P_\tau P_{\tau_1}MOe^{-A(\tau_1+\tau)}$ can be reconstructed by the user by time t.

□

Procedure. The steps that are required to calculate the delay-incorporating user-predictable subspace, $\mathcal{P}^*_{\mathcal{H}}$ are as follows:

- Obtain all the required matrices to determine $\mathcal{O}^*_{\mathcal{H}}$.
- Calculate $\delta_0 - \delta_2$ from (4.22).
- Obtain P_τ from (4.24).
- Determine the delay-incorporating user-predictable subspace from (4.20).

$\square$

Corollary 5. *A delay-incorporating user-predictable space is also delay-incorporating user-observable.*

Proof. As in the proof in Corollary 4, if is straight forward to show that with A being a full rank matrix

$$\begin{aligned}\mathcal{P}^*_{\mathcal{H}} &\triangleq \mathcal{R}((e^{-A(\tau+\tau_1)})^T(MO)^T P^T_{\tau_1} P^T_\tau) \\ &= \mathcal{R}((e^{-A\tau_1})^T(MO)^T P^T_{\tau_1} P^T_\tau) \\ &\subseteq \mathcal{R}((e^{-A\tau_1})^T(MO)^T P^T_{\tau_1}),\end{aligned}$$

which proves that $\mathcal{P}^*_{\mathcal{H}} \subseteq \mathcal{O}^*_{\mathcal{H}}$.

$\square$

Corollary 6. *A delay-incorporating user-predictable space is also user-predictable.*

Proof. Consider $\tau_1 = 0$, hence $P_{\tau_1} = I$ and from (4.20), $\mathcal{P}_{\mathcal{H}}$ can be formulated as

$$\mathcal{P}_{\mathcal{H}} \triangleq \mathcal{R}((e^{-A(\tau)})^T(MO)^T P^T_\tau). \tag{4.25}$$

As in Corollary 4 and 5, it is trivial to show that $\mathcal{P}^*_{\mathcal{H}} \subseteq \mathcal{P}_{\mathcal{H}}$

$\square$

4.3.1 Validation of the displayed information

In Section 4.1 we stated that for safety-critical systems under human or shared control, the user-interface must provide the user with information that results in a delay-incorporating user-observable and a delay-incorporating user-predictable task. Hence, we can introduce the following proposition.

Proposition 1. *In order for a user to be able to accomplish a desired task in a safety-critical condition for a system of form (2.1) and under the assumptions 7-4, the following inclusion is necessary*

$$\mathcal{T} \subseteq \mathcal{P}^*_{\mathcal{H}}, \tag{4.26}$$

where $\mathcal{P}^*_{\mathcal{H}}$ is the delay-incorporating user-predictable subspace, formulated in (4.20).

4.4 Examples

We consider a remotely driven point mass car modeled as a double integrator and stabilized to have poles on -2 and -3. The system matrices are

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{4.27}$$

Our goal is to evaluate whether for such a system the displays (which measure the position or the velocity of the car) are effective to accomplish a desired task, when the processing delay is $\tau_1 = 0.2$.

We consider two cases, 1) a user who controls the system via a known force with known constant rate – that is, all derivatives of the input are known and 2) a user whose input to the system is complicated and random, thus, has no knowledge about input derivatives – that is , $\lambda = 0$. For both cases, we consider $\gamma = 1$.

Our desired task is stopping at a stop sign, hence, we can define the task space as $\mathcal{T} = \mathcal{R}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)$ which spans $\mathbb{R}^2$.

4.4.1 delay-incorporating user-observable subspace

It is first required to calculate the delay-incorporating user-observable subspace for different measurements of states available to the user.

From (4.9), for $\lambda = \infty$, we can obtain $M = I$ and $P_{\tau_1} = I$. Hence,

$$\mathcal{O}^*_{\mathcal{H}}|_{\lambda=\infty} = \mathcal{R}(C^T) \oplus \mathcal{R}((e^{-A\tau_1})^T O^T). \tag{4.28}$$

Also, for $\lambda = 0$, we can obtain $M = I$ therefore,

$$\mathcal{O}^*_{\mathcal{H}}|_{\lambda=0} = \mathcal{R}(C^T) \oplus \mathcal{R}((e^{-A\tau_1})^T O^T P^T_{\tau_1}). \tag{4.29}$$

From (4.28) and (4.29), for either of the two different displays including a GPS with $C = [1\ 0]$ and a speedometer with $C = [0\ 1]$, we can show that $\mathcal{O}^*_{\mathcal{H}}|_{\lambda=\infty}$ spans $\mathbb{R}^2$.

In addition, when $\lambda = 0$, with either of the measurements in the display we can obtain

$$P_{\tau_1} = \begin{bmatrix} 0.9919 & 0.0899 \\ 0.0899 & 0.0081 \end{bmatrix},$$

hence, $\mathcal{O}^*_{\mathcal{H}}|_{\lambda=0} = \mathbb{R}^2$.

The results state that for such a system, regardless of the type of the measurements and the complexity of users' input, the user can reconstruct both states of the system.

4.4.2 Delay-incorporating user-predictable subspace

For $\tau = 0.1$, we calculate the delayed-incorporates user-predictable subspace of this system for different measurements of states provided in the display.

From (4.20), for $\lambda = \infty$, we can obtain $P_\tau = I$, hence

$$\mathcal{P}^*_{\mathcal{H}}|_{\lambda=\infty} = \mathcal{R}((e^{-A(\tau+\tau_1)})^T O^T) \tag{4.30}$$

Also, for $\lambda = 0$ and from (4.20)

$$\mathcal{P}^*_{\mathcal{H}}|_{\lambda=0} = \mathcal{R}((e^{-A(\tau+\tau_1)})^T O^T P^T_{\tau_1} P^T_{\tau}) \tag{4.31}$$

From (4.30), we can obtain $\mathcal{P}^*_{\mathcal{H}}|_{\lambda=\infty} = \mathbb{R}^2$ with either the GPS or the speedometer.

When $\lambda = 0$, for either of the displays,

$$P_\tau = \begin{bmatrix} 0.0081 & -0.0899 \\ -0.0899 & 0.9919 \end{bmatrix},$$

hence, $\mathcal{P}^*_{\mathcal{H}}|_{\lambda=0} = \emptyset$. This states that, with a complicated input, the user can reconstruct no combination of the states $x(t+\tau)$ at time t.

The results above can help the reader understand Corollary 5 better, as it is clear that in all of the above cases, the delay-incorporating user-predictable space is a subset of the delayed incorporated user-observable space.

We now consider $\lambda = 0$ for the case of having no processing delay, $\tau_1 = 0$. We thus can obtain the user-predictable subspace to be

$$\mathcal{P}_{\mathcal{H}} = \mathcal{R}\left(\begin{bmatrix} 1.0000 \\ -0.0526 \end{bmatrix}\right)$$

for either of the displays. Based on Corollary 6, the delay-incorporating user-predictable space is always a subset of the user-predictable space which we also have shown it to be the case in this example.

The result of having no delay shows that not considering the processing delay can result in a larger user-predictable subspace. Overlooking this delay can mislead the designer to misjudge the capability of the user to accomplish a task; i.e., by ignoring the delay, the designer might find the user capable of accomplishing the task, while, due to the existence of the processing delay, the space which is predictable by the user may not include the task space or even is empty. Thus, for safety critical systems, it is not safe to simply ignore this value as it may result in hazardous outcomes.

4.4.3 Task accomplishment

With any of the suggested displays, when $\lambda = \infty$, both the delay-incorporating user-observable and the delay-incorporating user-predictable spaces span $\mathbb{R}^2$. Hence, $\mathcal{T} \subset \mathcal{P}_{\mathcal{H}}*|_{\lambda=\infty}$. This means, regardless of the displayed information, if the pattern of changes of the user's input is all clear to the user, it might be possible for the user to stop at a stop sign.

On the other hand, it is clear that $\mathcal{T} \nsubseteq \mathcal{P}_{\mathcal{H}} * |_{\lambda=0}$ for either of the displays. Hence, under the Assumption 5, for a complicated input to the system, neither the GPS nor the speedometer are effective for a human to control the velocity of the mass. Hence, for such an input, regardless of the type of the measurements, there is always a chance that the user cannot estimate and predict the task precisely.

For this example, although not having a processing delay is helpful in expanding the user predictable space, it still does not help with task accomplishment.

4.5 Summary and conclusion

In this chapter, for safety-critical LTI systems, two novel subspaces, the delay-incorporating user-observable subspace, $\mathcal{O}^*_{\mathcal{H}}$, and the delay-incorporating user-predictable subspace $\mathcal{P}^*_{\mathcal{H}}$, with possibly longer term predictions, were formulated. As in Chapter 3, these subspaces were compared to the task space for a feasible task. If the task space does not lie in the relevant space, then the user-interface of a safety-critical system is incorrect, meaning that in such a system there exists a possibility that the user cannot accomplish the desired task with the given user-interface.

In the next two chapters, we suggest models for the process of attaining SA by the user. These models let us evaluate the correctness of a given display. For cases with unchanging operating conditions, with the aid of the mentioned models, the information which is required to be included in the display for the safety of the task can also be determined.

Chapter 5

User-interface analysis through modeling the user as an estimator

Having a system as in (2.1), our goal in this chapter is to introduce a detailed model of a human attaining SA.

In Sections 5.2.1 and 5.2.2, we present the existence conditions and the design procedure of an extended delayed functional observer/predictor considered to be the model of attaining SA by the user. We follow by an example on the existence and design of the delayed/non-delayed functional observer/predictor. In section 5.2.4, we suggest a technique to determine the required information to be displayed. Finally, in Section 5.3, we investigate a safety critical application, prediction of the depth of anesthesia during surgery.

5.1 Problem formulation

To model the process of attaining SA by the user, we take into account the users' limitations and capabilities regarding the information presented to them and estimated by them. The process of attaining SA includes observations as well as predictions by the user. Processing of information generally introduces a delay [104–106]. Assuming the derivatives of the inputs and the outputs might be available, we design (and evaluate the existence of) a novel estimator for LTI systems generating delayed estimates of the current and upcoming desired functional of states. Since we consider the user to i) only reconstruct and predict the desired set of states rather than the entire state space, ii) make delayed estimations, and iii) possibly have knowledge of the derivatives of the inputs and outputs, we model the process of attaining SA as an extended delayed functional observation/prediction.

Since, in this chapter, reconstruction and prediction of the desired states of the system are considered to be delayed, we model the user as a delayed observer/predictor for the functional:

$$z_0(t+\tau) = Tx(t+\tau), \qquad 0 \leq \tau, \tag{5.1}$$

where τ defines the prediction horizon. In (5.1), the task matrix $T \in \mathbb{R}^{l\times n}$ is defined in (2.3).

In some cases, it is not possible to estimate the functional $z_0(t+\tau)$ directly and it is necessary for the user to also estimate the functional $Rx(t+\tau)$ such that $R \in \mathbb{R}^{\mathcal{X}\times n}$. We select the rows of R to be linearly independent from the rows of T. Hence, we introduce the extended functional as

$$z(t+\tau) = \begin{bmatrix} T \\ R \end{bmatrix} x(t+\tau), \tag{5.2}$$

where R is selected such that $L = [T^T, R^T]^T$ is of full row rank. For cases that $Tx(t+\tau)$ can be estimated directly, we have $R = \emptyset$.

We introduce a theorem providing conditions needed for an estimator with human specifications to exist, i.e., the required conditions on the system, the display, and the task so that the user can attain SA toward specific goals. If these conditions are not satisfied for a triplet of dynamics, measurements, and task, then it is not possible for the user to attain SA regarding the specific task through the available information. Possible solutions to such a problem could be i) modifying the content of the display, and/or ii) expanding the task. In some systems, however, none of the above modifications help in attaining SA. This non-existence of the observer/predictor can itself be informative to the system designer, e.g., for making better decisions on how to provide the required information to the user.

The analyses in this chapter are under Assumption 3 for $\gamma \in \{0, 1\}$ and

$\lambda \in \{0,1\}$. It is worth mentioning that employing a technique similar to the one in this chapter, it is straightforward to design a non-delayed/delayed functional estimator for $1 \leq \gamma$ and $1 \leq \lambda$.

5.2 Methodology

Based on Assumption 3, we introduce the extended output vector and the extended input vector as

$$\begin{aligned} Y_{0:\gamma}(t) &= [{y_1}^T(t), \dot{y}_1^T(t), \cdots, {y_1^{(\gamma)}}^T(t)]^T, \\ U_{0:\lambda}(t) &= [{u_h}^T(t), \dot{u}_h^T(t), \cdots, {u_h^{(\lambda)}}^T(t)]^T, \end{aligned} \tag{5.3}$$

hence, as we consider $\gamma \in \{0,1\}$ and $\lambda \in \{0,1\}$, only two cases for the extended output vector and two cases for the extended input vector may exist. Analytically, the extended output vector can be written as

$$Y_{0:\gamma}(t) = O_\gamma x(t) + M_{1,0:\gamma} U_{0:\gamma}(t) + M_{2,\gamma} r_a, \tag{5.4}$$

where for $\gamma \in \{0,1\}$, the observability matrix $O_\gamma \in \mathbb{R}^{(\gamma+1)p_x \times n}$, a Toeplitz matrix $M_{1,0:\gamma} \in \mathbb{R}^{(\gamma+1)p_x \times (\gamma+1)p_x}$, and matrices $M_{2,\gamma} \in \mathbb{R}^{(\gamma+1)p_x \times p_r}$ and $U_{0:\gamma}(t) \in \mathbb{R}^{(\gamma+1)p_x}$ are defined as follows,

$$\begin{aligned} O_0 &= C \quad , & O_1 &= [C^T, A^T C^T]^T \\ M_{1,0:0} &= 0 \quad , & M_{1,0:1} &= \begin{bmatrix} 0 & 0 \\ CB & 0 \end{bmatrix} \\ M_{2,0} &= 0 \quad , & M_{2,1} &= \begin{bmatrix} 0 \\ CF \end{bmatrix} \\ U_{0:0}(t) &= u_h \quad , & U_{0:0}(t) &= [{u_h}^T(t), \dot{u}_h^T(t)]^T. \end{aligned} \tag{5.5}$$

In addition to the inputs, outputs, and their derivatives, we also give the user the ability to incorporate the measured trajectories in estimating the desired states. We therefore aim to model the user as an estimator of the form

$$\begin{aligned} \dot{\omega}(t) &= N\omega(t) + J_1 Y_{0:\gamma}(t) + J_2 y_2(t) + H U_{0:\lambda}(t) \\ \hat{z}(t) &= \omega(t - \tau_1) + E Y_{0:\gamma}(t) \end{aligned} \tag{5.6}$$

which produces delayed or non-delayed estimates of current or upcoming values of a desired functional of states. In (5.6), $\omega(t) \in \mathbb{R}^{l+\mathcal{X}}$ is the state of the estimator and τ_1 is the estimation delay. It is desirable to determine a stable matrix N and matrices J_1, J_2, H, and E with compatible dimensions to make the estimation error asymptotically approach zero. From (5.6), it is clear that we only apply the delay term on the desired states which need to be processed and estimated in the working memory – that is, the set of desired states which are not directly available to the user.

Having the estimator (5.6) to estimate the functional (5.2) of the system (2.1), the prediction error is

$$\begin{aligned} e(t) &= \hat{z}(t) - z(t+\tau) \\ &= \omega(t-\tau_1) + EY_{0:\gamma}(t) - Lx(t+\tau) \\ &= \omega(t-\tau_1) + EO_\gamma x(t) + EM_{1,0:\gamma}U_{0:\gamma}(t) + \\ &\quad\; EM_{2,\gamma}r_a - Lx(t+\tau) \end{aligned} \tag{5.7}$$

with the error dynamics

$$\begin{aligned} \dot{e}(t) &= N\omega(t-\tau_1) + J_1O_\gamma x(t-\tau_1) + J_1M_{1,0:\gamma}U_{0:\gamma}(t-\tau_1) + J_1M_{2,\gamma}r_a + \\ &\quad\; J_2Dr_a + HU_{0:\lambda}(t-\tau_1) + EO_\gamma Ax(t) + EO_\gamma Bu_h(t) + EO_\gamma Fr_a + \\ &\quad\; EM_{1,0:\gamma}U_{1:\gamma+1}(t) - LAx(t+\tau) - LBu_h(t+\tau) - LFr_a. \end{aligned} \tag{5.8}$$

Note that in (5.8), by setting $\tau = 0$ we obtain the error dynamics for the observation and by setting $\tau > 0$ we obtain the error dynamics for the prediction.

Since in general the delay and the value of prediction horizon are small, we can write

$$\begin{aligned} u_h(t) &= u_h(t-\tau_1) + \tau_1\dot{u}_h(t-\tau_1), \\ u_h(t+\tau) &= u_h(t-\tau_1) + (\tau+\tau_1)\dot{u}_h(t-\tau_1) + \tau\tau_1\ddot{u}(t-\tau_1). \end{aligned} \tag{5.9}$$

The states of the continuous time system (2.1) evolve as

$$\begin{aligned} x(t+\tau) &= e^{A\tau}x(t) + c, \\ x(t) &= e^{A\tau_1}x(t-\tau_1) + c_1, \end{aligned} \tag{5.10}$$

where

$$\begin{aligned} c &= \int_t^{t+\tau} e^{A(t+\tau-T)}[B \quad F] \begin{bmatrix} u_h(T) \\ r_a \end{bmatrix} dT, \\ c_1 &= \int_{t-\tau_1}^{t} e^{A(t-T)}[B \quad F] \begin{bmatrix} u_h(T) \\ r_a \end{bmatrix} dT. \end{aligned} \tag{5.11}$$

By introducing the variables

$$\begin{aligned} \delta_1 &\triangleq (e^{A\tau} - I)A^{-1}B, \\ \delta_2 &\triangleq ((e^{A\tau} - I)A^{-1} - \tau I)A^{-1}B, \\ \delta_3 &\triangleq (e^{A\tau} - I)A^{-1}F, \\ \theta_1 &\triangleq (e^{A\tau_1} - I)A^{-1}B, \\ \theta_2 &\triangleq ((e^{A\tau_1} - I)A^{-1} - \tau_1 I)A^{-1}B, \\ \theta_3 &\triangleq (e^{A\tau_1} - I)A^{-1}F, \\ \eta_1 &\triangleq e^{A(\tau+\tau_1)}, \\ \eta_2 &\triangleq \delta_3 + e^{A\tau}\theta_3, \\ \eta_3 &\triangleq \delta_1 + e^{A\tau}\theta_1, \\ \eta_4 &\triangleq \delta_2 + \tau_1\delta_1 + e^{A\tau}\theta_2, \\ \eta_5 &\triangleq \tau_1\delta_2, \end{aligned} \tag{5.12}$$

and from (5.9) and (5.10), we can write

$$\begin{aligned} x(t) &= e^{A\tau_1}x(t-\tau_1) + \theta_1 u(t-\tau_1) + \theta_2 \dot{u}(t-\tau_1) + \theta_3 r_a, \\ x(t+\tau) &= \eta_1 x(t-\tau_1) + \eta_2 u(t-\tau_1) + \eta_3 \dot{u}(t-\tau_1) + \eta_4 \ddot{u}(t-\tau_1) + \eta_5 r_a. \end{aligned} \tag{5.13}$$

Hence, under the assumption that γ and λ are selected from $\{0, 1\}$, the error dynamics can be written as

$$\begin{aligned}\dot{e}(t) \quad = \quad & Ne(t) + (NL\eta_1 - LA\eta_1 + [E \ J_1 \ K \ J_2 \ H]Q_{1,1})x(t-\tau_1) + \\ & (NL\eta_2 - LA\eta_2 + [E \ J_1 \ K \ J_2 \ H]Q_{1,2} - LF)r_a + \\ & (NL\eta_3 - LA\eta_3 + [E \ J_1 \ K \ J_2 \ H]Q_{1,3} - LB)u_h(t-\tau_1) + \\ & (NL\eta_4 - LA\eta_4 + [E \ J_1 \ K \ J_2 \ H]Q_{1,4} - LB(\tau+\tau_1))\dot{u}_h(t-\tau_1) + \\ & (NL\eta_5 - LA\eta_5 + [E \ J_1 \ K \ J_2 \ H]Q_{1,5} - LB\tau\tau_1)\ddot{u}_h(t-\tau_1),\end{aligned} \tag{5.14}$$

where $K \triangleq J_1 - NE$ and $Q_{1,i}$ are defined in (5.15) for $i \in \{1, \cdots, 5\}$.

$$\begin{aligned} Q_1 \quad \triangleq \quad & \left[\begin{array}{c|c|c|c|c} Q_{1,1} & Q_{1,2} & Q_{1,3} & Q_{1,4} & Q_{1,5} \end{array} \right] \\ \triangleq \quad & \left[\begin{array}{c:c:c:c:c} O_\gamma A e^{A\tau_1} & O_\gamma(A\theta_3 + F) & O_\gamma(A\theta_1 + B) & O_\gamma(A\theta_2 + \tau_1 B) + M_{1,\gamma} & \tau_1 M_{1,\gamma} \\ O_\gamma(I - e^{A\tau_1}) & -O_\gamma\theta_3 & -O_\gamma\theta_1 & -O_\gamma\theta_2 - \tau_1 M_{1,\gamma} & 0 \\ O_\gamma e^{A\tau_1} & O_\gamma\theta_3 + M_{2,\gamma} & O_\gamma\theta_1 + M_{1,\gamma} & O_\gamma\theta_2 + \tau_1 M_{1,\gamma} & 0 \\ 0 & D & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & \lambda I & 0 \end{array} \right] \end{aligned} \tag{5.15}$$

In (5.15), $M_{1,0} = 0_{p_x \times m_b}$ and $M_{1,1} = \begin{bmatrix} 0 \\ CB \end{bmatrix}$.

From (5.14), a τ_1-delayed estimator exists and can be designed to estimate the desired functional $z(t+\tau)$ if and only if there exists a set (E, J_1, N, J_2, H) , where $H \triangleq [H_a \ \ H_b]$, with a stable N to always satisfy

$$\begin{aligned} NL\eta_1 + [E \ J_1 \ K \ J_2 \ H]Q_{1,1} &= Q_{2,1}, \\ NL\eta_2 + [E \ J_1 \ K \ J_2 \ H]Q_{1,2} &= Q_{2,2}, \\ NL\eta_3 + [E \ J_1 \ K \ J_2 \ H]Q_{1,3} &= Q_{2,3}, \\ NL\eta_4 + [E \ J_1 \ K \ J_2 \ H]Q_{1,4} &= Q_{2,4}, \\ NL\eta_5 + [E \ J_1 \ K \ J_2 \ H]Q_{1,5} &= Q_{2,5}, \end{aligned} \tag{5.16}$$

with $Q_{2,i}$ are defined in (5.17) for $i \in \{1, \cdots, 5\}$.

$$\begin{aligned} Q_2 &\triangleq \left[\begin{array}{c|c|c|c|c} Q_{2,1} & Q_{2,2} & Q_{2,3} & Q_{2,4} & Q_{2,5} \end{array} \right] \\ &\triangleq \left[\begin{array}{c|c|c|c|c} LA\eta_1 & L(A\eta_2 + F) & L(A\eta_3 + B) & L(A\eta_4 + B(\tau + \tau_1)) & L(A\eta_5 + B\tau\tau_1) \end{array} \right] \end{aligned} \tag{5.17}$$

In summary, using the above conditions, we can formulate the problem as follows. We seek to:

- evaluate the satisfaction of (5.16) for a desired task T, a given delay τ_1, and a given amount of prediction horizon τ to determine whether it is possible for the user to attain SA regarding the desired task and thus make correct decisions toward its accomplishment.
- obtain the model of the user by solving (5.16) for N and $\begin{bmatrix} E & J_1 & K & J_2 & H \end{bmatrix}$ (which also satisfy $K = J_1 - NE$).
- seek the triplet (C, D, R) (if there exists any), with a minimum cardinality of (C,D), which satisfies conditions in (5.16) to determine the required information to be displayed. Note that, we define the cardinality of (C,D) as $rank(C) + rank(D)$.

5.2.1 Existence conditions for an extended functional estimator

For LTI systems under shared-control and assuming availability of the derivatives of the inputs and outputs, we obtain the necessary and sufficient conditions for the existence of a delayed/non-delayed functional observer and predictor.

Recall that we consider a full row rank functional $Lx(t+\tau)$, with $Lx(t+\tau) = \begin{bmatrix} T \\ R \end{bmatrix} x(t+\tau)$, whose components are $Tx(t+\tau)$ and $Rx(t+\tau)$. Therefore reconstructing $Lx(t+\tau)$ is sufficient for the reconstruction of the desired task, $Tx(t+\tau)$. Our goal is to investigate the existence of and then design an observer of form (5.6) to reconstruct the functional $Lx(t+\tau)$. Mathematically this is equivalent to finding a solution for (5.16).

Lemma 1. *There exists a solution for (5.16) iff the following two conditions are simultaneously satisfied:*

- $$\begin{bmatrix} E & J_1 & K & J_2 & H \end{bmatrix} T_1 = T_2, \tag{5.18}$$

  where $T_1 = Q_1 M_E$ and $T_2 = Q_2 M_E$ and M_E is from (5.20).

- $$\begin{aligned} N &= Q_{2,i}H_i - \begin{bmatrix} E & J_1 & K & J_2 & H \end{bmatrix} Q_{1,i}H_i, \\ & \quad for\ i \in \{1, \cdots, 5\}. \end{aligned} \tag{5.19}$$

  where H_i is such that $L\eta_i H_i = I$, for $i \in \{1, \cdots, 5\}$.

Proof. By selecting E_is to satisfy $L\eta_i E_i = 0$ and H_is defined earlier, we can define a full row-rank matrix

$$S_1 = \left[\begin{array}{c|c} M_H & M_E \end{array}\right], \tag{5.20}$$

where

$$M_E = \begin{bmatrix} E_1 & 0 & 0 & 0 & 0 \\ 0 & E_2 & 0 & 0 & 0 \\ 0 & 0 & E_3 & 0 & 0 \\ 0 & 0 & 0 & E_4 & 0 \\ 0 & 0 & 0 & 0 & E_5 \end{bmatrix}, \quad M_H = \begin{bmatrix} H_1 & 0 & 0 & 0 & 0 \\ 0 & H_2 & 0 & 0 & 0 \\ 0 & 0 & H_3 & 0 & 0 \\ 0 & 0 & 0 & H_4 & 0 \\ 0 & 0 & 0 & 0 & H_5 \end{bmatrix}. \tag{5.21}$$

Given that S_1 is of full row rank, post multiplication of S_1 in (5.16) will not change the results. As a result of this post multiplication, (5.18) and (5.19) are obtained to be an equivalent expression to (5.16). □

In order for a stable solution for (5.16) to exist, a stable matrix N and matrices $[E\ J_1\ K\ J_2\ H]$ have to exist to satisfy both (5.18) and (5.19).

Clearly, there exists a solution for (5.18) iff span$({T_2}^T) \subseteq \mathcal{R}({T_1}^T)$ – that is,

$$rank\begin{bmatrix} T_1 \\ T_2 \end{bmatrix} = rank\begin{bmatrix} T_1 \end{bmatrix}. \tag{5.22}$$

Proposition 2. *The condition (5.18) is satisfied iff*

$$rank(LHS_1) = rank(RHS) \tag{5.23}$$

where

$$RHS \triangleq \begin{bmatrix} & & Q_1 & & \\ L\eta_1 & L\eta_2 & L\eta_3 & L\eta_4 & L\eta_5 \end{bmatrix}, \tag{5.24}$$

and

$$LHS_1 \triangleq \begin{bmatrix} Q_2 \\ RHS \end{bmatrix}. \tag{5.25}$$

Proof. We can post-multiply S_1 from (5.20) in (5.24) and (5.25) to obtain

$$\begin{aligned} rank(RHS) &= rank(RHS \times S_1) \\ &= rank(L) + rank(T_1) \end{aligned} \tag{5.26}$$

and

$$\begin{aligned} rank(LHS_1) &= rank(LHS_1 \times S_1) \\ &= rank(L) + rank(\begin{bmatrix} T_1 \\ T_2 \end{bmatrix}) \end{aligned} \tag{5.27}$$

respectively. From (5.26) and (5.27), we can show that $rank(RHS) = rank(LHS_1)$ iff $rank\begin{bmatrix} T_1 \\ T_2 \end{bmatrix} = rank\begin{bmatrix} T_1 \end{bmatrix}$. Thus, (5.23) is the necessary and sufficient condition for the existence of the solution to (5.18). $\square$

Proposition 3. *The condition in (5.19) is satisfied, with a stable N, iff the following conditions are simultaneously satisfied.*

1. *For all $s \in \mathbb{C}$,*

$$rank(LHS_{2,i}) = rank(RHS) \tag{5.28}$$

where $LHS_{2,i}$ is formulated in (5.29). In (5.29), $M_{E,i:j}$ is a block diagonal portion of M_E, defined in (5.21), which only contains E_k on *its diagonal where* $k \in \{1, \cdots, j\}$*. In (5.28),* $i \in \{1, \cdots, 5\}$ *when* $\tau \neq 0$ *and* $\tau_1 \neq 0$*,* $i = 1$ *when* $\tau = 0$ *and* $\tau_1 = 0$*, and* $i \in \{1, \cdots, 4\}$ *when* $\tau = 0$ *and* $\tau_1 \neq 0$*.*

$$
\begin{aligned}
LHS_{2,1} &\triangleq \left[\begin{array}{ccccc} sL\eta_1 - LA\eta_1 & -Q_{2,2} & -Q_{2,3} & -Q_{2,4} & -Q_{2,5} \\ \hdashline \multicolumn{5}{c}{Q_1} \end{array}\right] C_{2,1}, \\
LHS_{2,2} &\triangleq \left[\begin{array}{ccccc} -Q_{2,1} & sL\eta_2 - LA\eta_2 & -Q_{2,3} & -Q_{2,4} & -Q_{2,5} \\ \hdashline \multicolumn{5}{c}{Q_1} \end{array}\right] C_{2,2},
\end{aligned}
$$

$$
\begin{aligned}
LHS_{2,3} &\triangleq \left[\begin{array}{ccccc} -Q_{2,1} & -Q_{2,2} & sL\eta_3 - LA\eta_3 & -Q_{2,4} & -Q_{2,5} \\ \hdashline \multicolumn{5}{c}{Q_1} \end{array}\right] C_{2,3}, \\
LHS_{2,4} &\triangleq \left[\begin{array}{ccccc} -Q_{2,1} & -Q_{2,2} & -Q_{2,3} & sL\eta_4 - LA\eta_4 & -Q_{2,5} \\ \hdashline \multicolumn{5}{c}{Q_1} \end{array}\right] C_{2,4}, \\
LHS_{2,5} &\triangleq \left[\begin{array}{ccccc} -Q_{2,1} & -Q_{2,2} & -Q_{2,3} & -Q_{2,4} & sL\eta_4 - LA\eta_4 \\ \hdashline \multicolumn{5}{c}{Q_1} \end{array}\right] C_{2,5}.
\end{aligned} \tag{5.29}
$$

where

$$
\begin{aligned}
C_{2,1} &\triangleq \begin{bmatrix} I & 0 \\ 0 & M_{E,2:5} \end{bmatrix}, & C_{2,2} &\triangleq \begin{bmatrix} E_1 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & M_{E,3:5} \end{bmatrix}, \\
C_{2,3} &\triangleq \begin{bmatrix} M_{E,1:2} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & M_{E,4:5} \end{bmatrix}, & C_{2,4} &\triangleq \begin{bmatrix} M_{E,1:3} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & E_5 \end{bmatrix}, \\
C_{2,5} &\triangleq \begin{bmatrix} M_{E,1:4} & 0 \\ 0 & I \end{bmatrix}.
\end{aligned}
$$

2. *When* $\tau_1 \neq 0$ *and/or* $\tau \neq 0$*, there exists a* Z *for which* $(\Lambda_1 - Z\Gamma_1)$ *has*

negative eigenvalues with acceptable magnitude

$$(\Lambda_1 - \Lambda_i) - Z(\Gamma_1 - \Gamma_i) = 0 \tag{5.30}$$

for $i \in \{2, \cdots, 5\}$ *when* $\tau_1 \neq 0$ *and for* $i \in \{2, \cdots, 4\}$ *when* $\tau = 0$.

In (5.30), Γ_i *and* Λ_i *are*

$$\begin{aligned} \Lambda_i &= Q_{2,i}H_i - T_2{T_1}^{+}Q_{1,i}H_i, \\ \Gamma_i &= (I - T_1{T_1}^{+})Q_{1,i}H_i. \end{aligned}$$

Proof. The solution to (5.18) has the form of

$$\begin{bmatrix} E & J_1 & K & J_2 & H \end{bmatrix} = T_2{T_1}^{+} + Z(I - T_1{T_1}^{+}) \tag{5.31}$$

for an arbitrary matrix Z with compatible dimension.

1. Proof of (5.28).

   From (5.19) and (5.31), N can be written as $N = \Lambda_i - Z\Gamma_i$ where

   $$\begin{aligned} \Lambda_i &= Q_{2,i}H_i - T_2{T_1}^{+}Q_{1,i}H_i, \\ \Gamma_i &= (I - T_1{T_1}^{+})Q_{1,i}H_i. \end{aligned} \tag{5.32}$$

   The eigenvalues of N can be placed at any desired values iff $rank \begin{bmatrix} sI - \Lambda_i \\ \Gamma_i \end{bmatrix} = l + \mathcal{X},\ \forall s \in \mathbb{C}$.

   We first introduce some required matrices to complete the proof. Choose full-row rank matrices

$$
\begin{aligned}
S_{a,1} &= \left[\begin{array}{cc|c} H_1 & E_1 & 0 \\ \hline 0 & 0 & I \end{array}\right], \\
S_{a,i} &= \left[\begin{array}{cccc} 0 & I & 0 & 0 \\ \hline H_i & 0 & E_i & 0 \\ \hline 0 & 0 & 0 & I \end{array}\right], \quad for\ i \in \{2,3,4\} \\
S_{a,5} &= \left[\begin{array}{cc|c} 0 & 0 & I \\ \hline H_1 & E_1 & 0 \end{array}\right],
\end{aligned} \tag{5.33}
$$

a full column rank matrix

$$
S_b = \begin{bmatrix} I & T_2{T_1}^+ \\ 0 & (I - T_1{T_1}^+) \\ 0 & T_1{T_1}^+ \end{bmatrix}, \tag{5.34}
$$

and a full row rank matrix

$$
S_{c,i} = \begin{bmatrix} I & 0 \\ {T_1}^+ Q_{1,i} H_i & I \end{bmatrix} \quad for\ i \in \{1, \cdots, 5\}. \tag{5.35}
$$

For each feasible value of i, by first post-multiplying $S_{a,i}$ in $LHS_{2,i}$ from (5.29) and then pre-multiplying S_b and post-multiplying $S_{c,i}$ in the resulted matrix, the rank does not change and we have

$$
rank(LHS_{2,i}) = rank \begin{bmatrix} sI - \Lambda_i \\ \Gamma_i \end{bmatrix} + rank[T_1]. \tag{5.36}
$$

Comparing (5.36) and (5.26) and having $rank(L) = l + \mathcal{X}$, (5.28) is satisfied iff $rank \begin{bmatrix} sI - \Lambda_i \\ \Gamma_i \end{bmatrix} = l + \mathcal{X},\ \forall s \in \mathbb{C}$.

2. Proof of (5.30).

   In proof of (5.28) we showed that, for each feasible value of i, the eigenvalues of matrix N can be selected to have a desired values depending on a matrix Z if $rank(LHS_{2,i}) = rank(RHS)$.

After satisfaction of (5.28), according to $N = \Lambda_i - Z\Gamma_i$, required is a common pair (N, Z) with an stable N which can satisfy the above equation for all feasible values of i. Thus, for any feasible pair of (i, j), it is necessary to have $\Lambda_i - Z\Gamma_i = \Lambda_j - Z\Gamma_j$, which is the proof to (5.30).

□

From Lemma 1 and Propositions 2-3, we can introduce Theorem 5 on the existence of a stable delayed functional estimator to estimate the current and upcoming desired functional of the states of a system of interest.

Theorem 5. *For a system of the form (2.1), with $\gamma \in \{0, 1\}$ available derivatives of the outputs and $\lambda \in \{0, 1\}$ available derivatives of the inputs, there exists an estimator of the form (5.6) to make*

- *non-delayed observations, with $\tau_1 = 0$ and $\tau = 0$,*
- *delayed observations, with $\tau_1 \neq 0$ and $\tau = 0$,*
- *non-delayed predictions, with $\tau_1 = 0$ and $\tau \neq 0$,*
- *delayed predictions, with $\tau_1 \neq 0$ and $\tau \neq 0$,*

iff, a functional Rx exists to extend the desired task functional; as is defined in (5.2); to satisfy the condition (5.23) in Proposition 2 and conditions (5.28) and (5.30) in Proposition 3 and also to satisfy the following condition:

- *When $\tau_1 \neq 0$, there exists a pair (N, Z), with Z satisfying (5.30) and N being a stable matrix, to hold*

$$N(T_{a,1} + ZT_{b,1}) = (T_{a,2} - T_{a,3}) + Z(T_{b,2} - T_{b,3}), \tag{5.37}$$

where

$$\begin{aligned} \begin{bmatrix} T_{a,1} & T_{a,2} & T_{a,3} & T_{a,4} \end{bmatrix} &\triangleq T_2 {T_1}^+, \\ \begin{bmatrix} T_{b,1} & T_{b,2} & T_{b,3} & T_{b,4} \end{bmatrix} &\triangleq (I - T_1 {T_1}^+), \end{aligned} \tag{5.38}$$

and have compatible dimensions ($T_{a,4}$ and $T_{b,4}$ have $(p_r + (\lambda+1)B)$ columns, and $T_{a,i}$ and $T_{b,i}$ have equal number of columns for $i \in \{1,2,3\}$).

In (5.38), $T_1 = Q_1 M_E$ and $T_2 = Q_2 M_E$ where Q_1 and Q_2 are obtained from (5.15) and (5.17). In (5.38), ${T_1}^+$ is the pseudo-inverse of T_1.

Proof. The necessary and sufficient condition for the existence of a delayed/non-delayed estimator of form (5.6) to reconstruct the functional $Tx(t+\tau)$ of the system (2.1) is the existence of a matrix R and a stable matrix N to satisfy equations (5.16). This problem can be considered as two subproblems: i) existence of a stable solution for (5.16), and ii) satisfaction of the condition $J_1 = K + NE$.

In the proofs of Propositions 2 and 3, we have already showed that the satisfaction of (5.23) is necessary and sufficient for the existence of solution to (5.18) and that satisfaction of (5.28) and (5.30) are necessary and sufficient for the existence of the solution to (5.19).

In addition to the existence of the solution for (5.18) and (5.19) and also the stability of matrix N, it is required to choose N to satisfy $J_1 \triangleq K + NE$.

When $\tau_1 = 0$, from (5.12), $\theta_i = 0$ which will result in $Q_{1,2} = 0$. Thus, J_1 can be selected arbitrarily and it will be straight forward to obtain N to satisfy $J_1 \triangleq K + NE$.

On the other hand, when $\tau_1 \neq 0$, J_1 cannot be selected arbitrarily. Therefore, we need to select the pair (N, Z) such that N is stable and $J_1 = K + NE$. Recall that, from (5.31), the selection of Z will affect the values of $[E\ J_1\ K\ J_2\ H]$. Equation (5.37) is obtained by plugging in (5.31) to $J_1 = K + NE$.

□

5.2.2 Model of the user as an estimator

Assuming the existence of the estimator (5.6) to reconstruct the functional $Tx(t+\tau)$ of the system (2.1), we can determine the related estimator matrices N, J_1, J_2, H, and E for delayed/non-delayed observation/prediction as follows:

- To design a non-delayed observer (i.e. for $\tau = 0$ and $\tau_1 = 0$),
  - Find a matrix R to satisfy (5.23) and (5.28), then form the matrix L.
  - Choose H_1 and E_1 such that $LH_1 = I$ and $LE_1 = 0$ respectively.
  - Calculate Λ_1 and Γ_1 from (5.32).
  - Choose Z to have a stable $N = \Lambda_1 - Z\Gamma_1$ and obtain J_2, E, K, and H from (5.31).
  - Based on N, E and K, calculate J_1 from $J_1 = K + NE$.
- To design a delayed observer (i.e. for $\tau = 0$ and $\tau_1 \neq 0$),
  - Find matrices R, N, and Z to simultaneously satisfy (5.23), (5.28), (5.30), and (5.37), then form the matrix L.
  - Based on Z, obtain J_1, J_2, E, K, and H from (5.31).
- To design a non-delayed predictor (i.e. for $\tau \neq 0$ and $\tau_1 = 0$),
  - Find matrices R, N, and Z to simultaneously satisfy (5.23), (5.28), and (5.30), then form the matrix L.
  - Determine J_2, E, K, and H from (5.31).
  - Based on N, E and K, calculate J_1 from $J_1 = K + NE$.
- To design a delayed predictor (i.e. for $\tau \neq 0$ and $\tau_1 \neq 0$),
  - Find matrices R, N, and Z to simultaneously satisfy (5.23), (5.28), (5.30), and (5.37), then form the matrix L.
  - Based on Z, obtain J_1, J_2, E,K, and H from (5.31).

5.2.3 Example

In this example, we validate our method for designing a delayed/non-delayed observer/predictor to estimate a desired functional. We assume $\gamma = 1$ (i.e., having access to the first derivative of the outputs) and $\lambda = 0$ (i.e., no derivative of the low-level input). Our goal is to evaluate the existence of a delayed-observer and a delayed-predictor to reconstruct $Tx(t+\tau)$, where $T = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}^T$ and then design such an estimator.

Consider a system of the form (2.1) with the following system matrices,

$$\begin{aligned} A &= \begin{bmatrix} 0 & 1 & -10 \\ -2 & -3 & -1 \\ 2 & 0 & -2 \end{bmatrix}, \\ B &= \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T, \\ F &= \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T, \end{aligned} \tag{5.39}$$

and $C = I_{3\times 3}$ – that is, all states are measured. Besides, the reference trajectory is assumed to be available in the user interface, $D = 1$. Given the system dynamics in (5.39), with all states being measured, the system is observable and predictable. However, based on our earlier discussion, the standard observability and predictability of the system is not enough for a human operator to accomplish the desired task. Through the conditions of Theorem 5, we can evaluate whether the human can attain SA about the task Tx.

For the system (5.39), the conditions in Theorem 5 for the existence of the delayed and non-delayed estimator where $\tau \in \{0, 0.2\}$ and $\tau_1 \in \{0, 0.3\}$ are satisfied. Therefore, based on the technique suggested in Section 5.2.2, we can design delayed/non-delayed observer/predictor for this system to reconstruct $Tx(t+\tau)$.

- Non-delayed observer ($\tau = 0$ and $\tau_1 = 0$):

  Using the design procedure suggested in Section 5.2.2, we can design a non-delayed observer as is illustrated in (5.40).

$$\begin{bmatrix} N \\ J_2 \\ H \end{bmatrix} = \begin{bmatrix} -4.2204 \\ 0.0546 \\ 1.2624 \end{bmatrix}, \qquad J_1 = \begin{bmatrix} -0.8084 \\ -0.2323 \\ -4.5347 \\ -0.1637 \\ -0.9632 \\ 0.3564 \end{bmatrix}^T,$$
$$E = \begin{bmatrix} 0.2515 \\ 0.2605 \\ 1.0778 \\ 0.1342 \\ 0.3544 \\ -0.0000 \end{bmatrix}^T, \qquad K = \begin{bmatrix} 0.2530 \\ 0.8672 \\ 0.0139 \\ 0.4028 \\ 0.5326 \\ 0.3564 \end{bmatrix}^T. \tag{5.40}$$

Figure 5.1 shows the effectiveness of using the designed observer in tracking the desired functional of the states of system (5.39) while the observer has no delay.

We now use the designed observer (same matrices as above) to predict our desired functional while the actual observer is delayed ($\tau_1 = 0.3$ and $\tau = 0.2$). The results are available in Figure 5.2.

From Figure 5.2, it can be seen that the non-delayed observer matrices are not effective for precisely predicting a desired functional while the structure of the actual observer is delayed too. Hence, a new estimator have to be designed to provide us with our desired results.

- Delayed-predictor ($\tau = 0.2$ and $\tau_1 = 0.3$):

  From the algorithm provided in Section 5.2.2, we can design a delayed predictor for the reconstruction of the functional Tx. The predictor

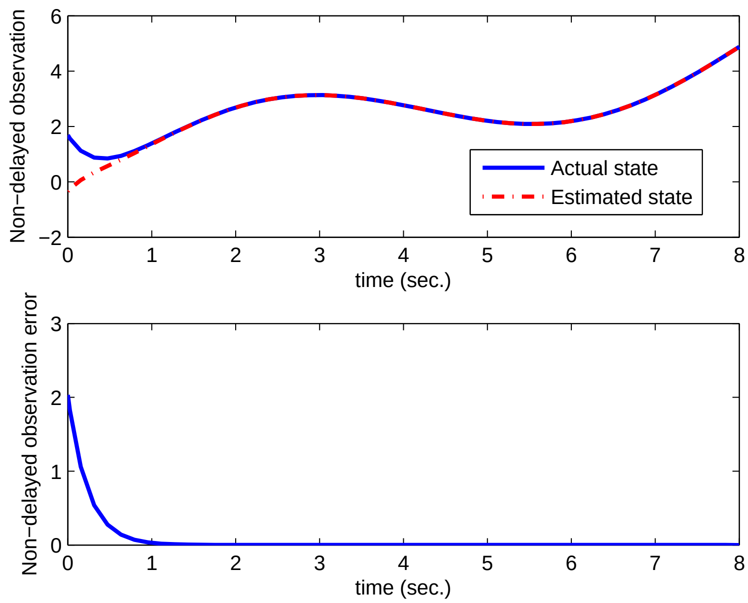


Figure 5.1: Non-delayed observation of the desired functional, Tx, of the states of the system (5.39).

structure is provided in 5.41.

$$\begin{bmatrix} N \\ J_2 \\ H \end{bmatrix} = \begin{bmatrix} -8.4719 \\ 0.1800 \\ 1.8711 \end{bmatrix}, \quad J_1 = \begin{bmatrix} 0.6541 \\ 0.2090 \\ 0.3986 \\ -0.1792 \\ -0.4676 \\ -1.1273 \end{bmatrix}^T, \tag{5.41}$$

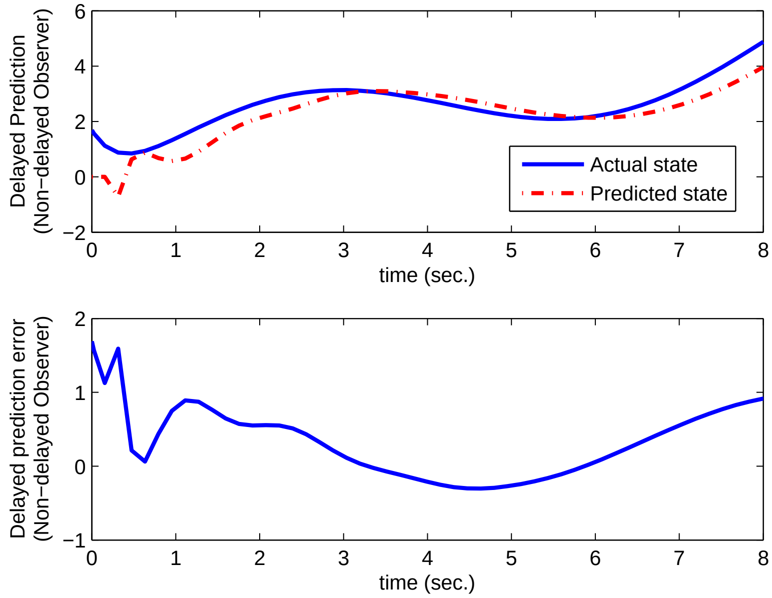


Figure 5.2: Using a non-delayed observer matrices for predicting Tx, while the actual observer has also internal delay.

$$E = \begin{bmatrix} -0.2100 \\ 0.1348 \\ 0.0114 \\ -0.0730 \\ 0.1040 \\ 0.2905 \end{bmatrix}^T, K = \begin{bmatrix} -1.1250 \\ 1.3510 \\ 0.4956 \\ -0.7978 \\ 0.4136 \\ 1.3335 \end{bmatrix}^T.$$

Figure 5.3 illustrates the simulation results of such a predictor. From this figure, it is possible to design a delayed estimator for system (5.39) to predict the desired functional, Tx, which confirms the results of Theorem 5.

Since the prediction starts when $t \geq \tau$, it will result in the discontinuity observed in Figure 5.3.

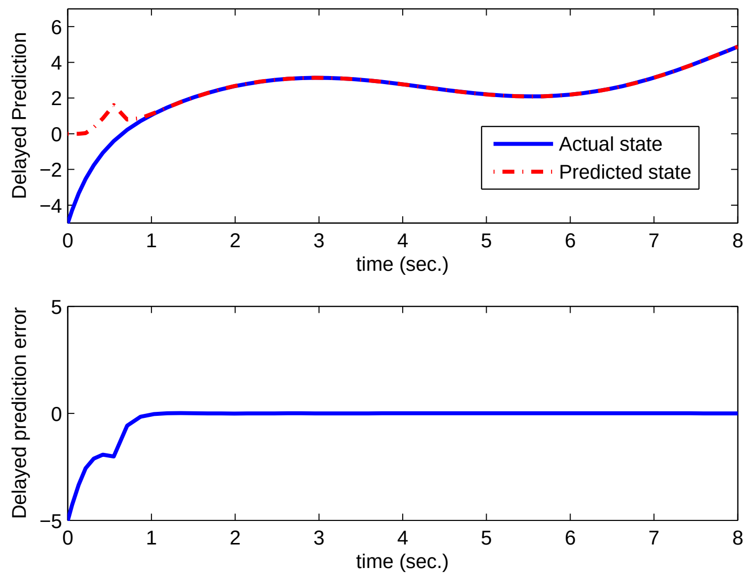


Figure 5.3: Delayed prediction of the desired functional, Tx, of the states of the system (5.39), with $\tau = 0.2$ and $\tau_1 = 0.3$.

5.2.4 Towards display design

In Section 5.2.1, we assumed the availability of certain measurements of the states and also the automation's desired trajectory and evaluated whether the human can make delayed/non-delayed estimations based on the available data. We also showed how the user can obtain the desired estimations.

Our main goal in this section is to discuss the stages required for designing a display with minimal cardinality to allow the user to accomplish a desired task. For this purpose, we have to find matrices C and D of minimum rank summation to satisfy the conditions in Theorem 5, which are required for the existence of a delayed predictor as well as a delayed observer (i.e., for prediction, $\tau = \tau_a$ and $\tau_1 = \tau_{1,a}$ with τ_a the prediction horizon amount and $\tau_{1,a}$ the information processing delay; and for observation, $\tau = 0$ and $\tau_1 = \tau_{1,a}$).

Clearly, desired for most designers is the *real-time determination of the* *displayed information* such that depending on the operating condition and the task, the display information can be updated. Analytically, our goal is to solve the problem

$$\begin{aligned} \min_{C,D} \quad & rank(C) + rank(D) \\ & \quad\quad subject\ to \quad\quad (5.23)\ for\ \tau = 0\ \&\ \tau = \tau_a, \\ & \quad\quad\quad\quad\quad\quad\quad\quad\ (5.28),\ (5.30),\ and\ (5.37). \end{aligned} \tag{5.42}$$

Remark. Since some measurements might be easier than others for the user to perceive and process, in some cases, the cardinality of the displayed information is not as important as the nature of the displayed information. Hence, in practice, it might be preferable for a designer to investigate several valid designs with low cardinalities and select the most suitable one. In addition, as the rank of matrix R is directly related to the required order of the functional estimator, having this matrix with low rank is also very important. Therefore, the reader may even consider selecting a feasible pair of (C, D) with their corresponding matrix R obtained to be of low or even minimum rank – that is, rather than (5.42), solving an optimization problem with $rank(R)$ being its cost function and the objective function as in (5.42). □

Although there are techniques to formulate the rank condition in a convex form, with the sophisticated constraints in (5.42), solving the mentioned problem and designing a display of minimum information for a generic case of online determination of the information is very complicated topic that should be a subject to extensive future research.

On the other hand, for cases where the operating conditions do not change and the task is pre-specified, the real-time determination of the display is not necessary and it is possible to simply design the displayed information in advance and apply it as is during the running of the process. We call this technique the *off-line determination of the displayed information*. For such simpler cases, it is possible to determine the required displayed information by manipulating the matrices C and D and verifying the satisfaction of the conditions in Theorem 5, heuristically.

For our specific application of user-interface design, C represents a set of the measured states and D represents the set of the automation's desired trajectories available on the user interface. Thus, we can consider the matrices C and D to be diagonal with the elements on the diagonal being zero or one – that is, no linear combinations of states and no linear combinations of the trajectories is presented in the interface. Clearly, if a specific linear combination of the states has a physical meaning to the user, that vector can also be added to the set of feasible measurements (e.g. the flight path angle of an aircraft dynamics which is a linear combination of the states pitch angle and the angle of attack).

Considering the processing delay, our goal is to determine the displayed information with minimum cardinality that lets us observe/predict the task functional. The required steps to determine the correct displayed information are as follows:

- Initialize by considering certain available displayed information, e.g, for initialization we suggest to have $rank(C) = 1$ and $rank(D) = 0$.
- For the available display, check the conditions in Theorem 5 (for $\tau_1 = 0$ and $\tau_1 = \tau_{1,a}$).
- If either the delayed functional observer or the delayed functional predictor does not exist, change the display information or iteratively increase the number of the measurements (including the measured states and the measured trajectories) and re-investigate the conditions in Theorem 5.
- Among the valid displays which satisfy all the existence conditions, select those of interest, either for lower rank, or for any application-specific reason which may be important to the designer.

As has already been mentioned, for most of the applications, the suggested method is necessary but not sufficient to design a good display. Hence, in order to design a display that is compatible with any operating condition and task, further research is required.

5.3 Application example

As for the users of any system with a human controller or under human supervision/monitoring, attaining SA is indispensable for an anesthetist to maintain the safety of the anesthetized patient. This importance has been investigated by several researchers [31, 107, 108]. In a recent paper, Fioratou et al. [107] discussed SA in the framework of anesthesiology. According to [107], after perceiving the available displayed information and the information from the environment, the anesthetist has to integrate all the available data for the identification of the current and the future desired patient states. The estimation of the current states of the system is important for goal accomplishment and for fault detection. As is mentioned in [107], task prediction is also extremely important for the anesthetist to be proactive rather than just being reactive.

In order to model a patient under anesthesia, understanding the relationship between the dose of the drug and its pharmacological effect is necessary. This model consists of two sub-models, the pharmacokinetic (PK) and the pharmacodynamic (PD) models. The PK model, demonstrates the effect of the administered drug on the drug plasma concentration and the PD model, models the relationship between the drug concentration in the effect site and the observed effect of the drug.

We consider a simplified version of the PKPD model described in [109] and [110] to model the effect of propofol administration on the depth of hypnosis. The PK model in [110] is the well-known 3-compartment model developed in [111] to evaluate the effect of propofol on the drug concentration in different compartments. Pharmacokinetically, a compartment is considered to be a group of tissues which have similar kinetic characteristics. A 3-compartment model has three states, i.e. concentrations in i) the blood and highly perfused tissues (e.g. brain and liver), ii) the muscles and viscera, and iii) fat and bones. The PD model presented in [110] consists of three states, two of which are associated with the dynamics of the monitor. We, however, only consider the effect site concentration of the drug and linearize the Hill equation to obtain the depth of hypnosis based on this state of the

k_{10}	0.0524	k_{12}	0.2359	k_{13}	0.0162
k_{21}	0.0892	k_{31}	0.0022	k_d	0.1335
V_1	0.3593	EC_{50}	3.2	γ_h	4.7

Table 5.1: PKPD coefficient for a 21 years old 100 kg patient

system.

Considering no transport delay, the PKPD model for evaluating the depth of anesthesia is as (2.1) with

$$\begin{aligned} A &= \begin{bmatrix} \multicolumn{3}{c}{A_{pk}} & 0 \\ k_d & 0 & 0 & -k_d \end{bmatrix}, \\ B &= \begin{bmatrix} B_{pk} \\ 0 \end{bmatrix}, \quad F = \begin{bmatrix} 0_{4\times 1} \end{bmatrix}, \end{aligned} \tag{5.43}$$

where

$$\begin{aligned} A_{pk} &= \begin{bmatrix} -(k_{10}+k_{12}+k_{13}) & k_{12} & k_{13} \\ k_{21} & -k_{21} & 0 \\ k_{31} & 0 & -k_{31} \end{bmatrix}, \\ B_{pk} &= \begin{bmatrix} V_1^{-1} \\ 0 \\ 0 \end{bmatrix}. \end{aligned} \tag{5.44}$$

In (5.43) and (5.44), k_{ij} and k_d are rate constants and V_1 is the volume of the plasma compartment.

The desired task which is controlling the depth of anesthesia can be defined as $Tx = \begin{bmatrix} 0 & 0 & 0 & \gamma_h(4EC_{50})^{-1} \end{bmatrix} x$, where EC_{50} is the 50% effect concentration and γ_h is the cooperativity coefficient. The values that we are using are presented in Table 5.1.

From Theorem 5, no delayed predictor for system (5.43) can reconstruct the desired functional $Tx(t+\tau)$ when the measurement in the display is restricted to the depth of hypnosis, $C = \begin{bmatrix} 0 & 0 & 0 & \gamma_h(4EC_{50})^{-1} \end{bmatrix}$. However, based on the discussion in Section 5.2.4, it can be seen that for an estimator of the form (5.6) with $\tau_1 = 0.3sec$ estimation delay, it is necessary and sufficient to measure the blood-plasma drug concentration in addition to the depth of hypnosis in order to make correct and precise observations and predictions of the desired functional – that is, $C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \gamma_h (4EC_{50})^{-1} \end{bmatrix}$.

Hence, from the above analysis we can conclude that it is not possible for the anesthetist to precisely predict the effect of the drug administration on the depth of hypnosis unless they are provided with both the depth of hypnosis and the plasma concentration through the display. Unfortunately, plasma concentration measurement is beyond current state of technology and thus cannot be provided. In practice, open-loop population-based PKPD models are used to estimate and display both C_p and C_e. However, due to significant inter-patient variability, these estimates come with such large uncertainties that they are not likely to substitute for real measurements, hence hindering the ability of the anesthetist to accurately predict the depth of hypnosis.

5.4 Summary and discussion

The focus of this chapter was on mathematical modeling of the process of attaining SA for the user. The user was considered to be a functional observer and a functional predictor whose estimations of the states of the system are delayed. It was also assumed that the user may have knowledge about the derivatives of their own inputs and of the outputs.

For a system that is controlled by the user and that tracks a desired reference trajectory with the aid of a computer, we presented a technique to evaluate whether it is possible to reconstruct and predict (with delay) a desired set of states given a set of displayed measurements. In addition to obtaining the existence conditions for the extended delayed functional observer/predictor with the availability of higher-order derivatives of the inputs and the outputs, we presented a procedure to design such an estimator. We also presented a method to determine the minimum information required to display so that the user can accomplish a desired task.

The work of this chapter focused on precise comprehension and prediction of the task by the user. In most of the systems, however, the users do not require to precisely comprehend and predict the task, but, it is enough, yet necessary, for them to make these estimations within a specific bound. That is the inspiration for the next chapter in which the process of attaining SA by the user will be modeled as a bounded-error delayed functional estimator.

Chapter 6

User-interface analysis through modeling the user as an estimator with bounded error

In this chapter, for a system modeled in (2.1), we model the process of attaining SA by the user as a bounded-error delayed functional estimator.

In Section 6.2 we will introduce a Theorem and Corollary on the existence of and the type of estimator that we are looking for. In 6.3 we talk about the anesthesia example in detail and use our tool to make some analysis on such systems.

6.1 Problem formulation

In the previous chapter, we investigated the correctness of and then designed the displayed information for safety critical systems in which for attaining SA, the precise comprehension and prediction of the information is necessary for their user. The previous results, however, were too restrictive for the majority of the systems in the real world. In most of the systems, the users do not require to precisely comprehend and predict the task, but, it is enough, yet necessary, for them to make these estimations within a specific bound. Consider a driver trying to maintain the speed of a car within the speed limits. This speed limit prevents the user from exceeding a specific speed while the user should also not drive too slowly. Hence, it is important for the driver to be capable of keeping the speed within the pre-specified bounds.

To model the user as a bounded-error delayed observer/predictor, we obtain the error dynamics as in Chapter 5. We then evaluate the conditions under which the error remains bounded for all feasible situations.

6.2 Bounded-error estimator

Having the error dynamics (5.8), our goal is to design the estimator matrices in (5.6) such that the steady-state error remains bounded for all feasible combinations of the inputs and the initial states.

Under the assumption that γ and λ are selected from $\{0, 1\}$, the error dynamics can be written as

$$\dot{e}(t) = Ne(t) + (NL\eta + [E\ J_1\ K\ J_2\ H]Q_1 - Q_2) \begin{bmatrix} x(t-\tau_1) \\ r_a \\ u_h(t-\tau_1) \\ \dot{u}_h(t-\tau_1) \\ \ddot{u}_h(t-\tau_1) \end{bmatrix}, \tag{6.1}$$

where $\eta = [\eta_1\ \eta_2\ \eta_3\ \eta_4\ \eta_5]$, $K \triangleq J_1 - NE$ and for $i \in \{1, \cdots, 5\}$, $Q_{1,i}$ and $Q_{2,i}$ are defined in (5.15) and (5.17) respectively.

We define a matrix $XU \in \mathbb{R}^{n \times i}$ which columns are selected to be all i feasible combinations of states, inputs, and input derivatives of a system, and formulate

$$\begin{aligned} C_1 &\triangleq \bar{e}_1 + L\eta XU \\ C_2 &\triangleq \bar{e}_2 + L\eta XU \\ T_1 &\triangleq Q_1 XU \\ T_2 &\triangleq Q_2 XU. \end{aligned}$$

with $\bar{e}_1 = \{e_1, e_1, \cdots e_1\}$ and $\bar{e}_2 = \{e_2, e_2, \cdots e_2\}$ have i columns.

Proposition 4. *For a system of form (2.1) with a given feasible combinations of state-input, XU, there exists a bounded error estimator of form*

(5.6) to reconstruct the functional $z(t+\tau)$ iff there exists a random vector V, a stable N, and a matrix Z, to satisfy

- $$T2 - NC_2 \leq VT_1 \leq T2 - NC_1. \tag{6.2}$$
- *when $\tau_1 \neq 0$,*

  $$N(ZT_{a,1} + VT_{b,1}) + Z(T_{a,2} - T_{a,3}) + V(T_{b,2} - T_{b,3}) = 0 \tag{6.3}$$

  where

  $$\begin{aligned} [T_{a,1}\ T_{a,2}\ T_{a,3}\ T_{a,4}] &\triangleq (I - T_1 {T_1}^{+}), \\ [T_{b,1}\ T_{b,2}\ T_{b,3}\ T_{b,4}] &\triangleq T_1 {T_1}^{+} \end{aligned}$$

  and $T_{a,i}$ and $T_{b,i}$ are of compatible dimensions.

Proof. We need to find the estimator matrices to always keep the steady state error bounded within the desired values. With the error dynamics provided in (5.8) and since we consider the state-input vector to be constant (a discussion will be provided later in this section) the error evolves as

$$e(t) = e^{Nt} e(0) + F_r \tag{6.4}$$

where the forced response, F_r, is

$$F_r = (e^{Nt} - I)N^{-1}(NL\eta + [E\ J_1\ K\ J_2\ H]Q_1 - Q_2) \begin{bmatrix} x(t-\tau_1) \\ r_a \\ u_h(t-\tau_1) \\ \dot{u}_h(t-\tau_1) \\ \ddot{u}_h(t-\tau_1) \end{bmatrix}.$$

Mathematically, we want the steady-state error to be bounded within pre-specified values e_1 and e_2 all the time. Considering that N will be designed to be stable, the boundedness can be formulated as

$$Ne_1 \leq (Q_2 - [E\ J_1\ K\ J_2\ H]Q_1 - NL\eta) \begin{bmatrix} x(t-\tau_1) \\ r_a \\ u_h(t-\tau_1) \\ \dot{u}_h(t-\tau_1) \\ \ddot{u}_h(t-\tau_1) \end{bmatrix} \leq Ne_2, \tag{6.5}$$

where $\leq$ shows the element-wise inequality.

From (6.5) we can write

$$NC_1 \leq C_T \leq NC_2 \tag{6.6}$$

where

$$C_T \triangleq T_2 - [E\ J_1\ K\ J_2\ H]T_1. \tag{6.7}$$

- **Proof of (6.2):** As $[E\ J_1\ K\ J_2\ H]T_1 = T_2 - C_T$, the solution exists for $[E\ J_1\ K\ J_2\ H]$ if and only if

$$rank[T_1] = rank \begin{bmatrix} T_2 - C_T \\ T_1 \end{bmatrix}, \tag{6.8}$$

which is equivalent to saying that for the existence of a solution to $[E\ J_1\ K\ J_2\ H]$ it is necessary and sufficient that there exists a C_T such that $T_2 - C_T$ be a linear combination of the rows of T_1. Hence, $[E\ J_1\ K\ J_2\ H]$ has infinite number of solutions for arbitrary values of V where

$$C_T = T_2 - VT_1. \tag{6.9}$$

From (6.6) and (6.9), Condition (6.2) is proved and for the stability of the observer, a stable N have to exist.

- **Proof of (6.3):** In addition to satisfaction of (6.2) and stability of matrix N, it is required to choose N to satisfy $J_1 \triangleq K + NE$.

  When $\tau_1 = 0$, J_1 can be selected arbitrarily. However, when $\tau_1 \neq 0$, J_1 is not an arbitrary matrix and we need to select the triplet (N, Z, V) such that N is stable and $J_1 = K+NE$. The solution to $[E\ J_1\ K\ J_2\ H]$ is as follows

$$\begin{aligned} [E\ J_1\ K\ J_2\ H] &= (T_2 - C_T)T_1{}^{+} + Z(I - T_1T_1{}^{+}) \\ &= Z(I - T_1T_1{}^{+}) + VT_1T_1{}^{+}. \end{aligned} \tag{6.10}$$

  From (6.10) and the condition $J_1 = K + NE$, (6.3) can be achieved.

$\square$

The steady-state error of the estimator will be maximized when the forced response of the error is maximum. This happens at a specific combination of N, $[E\ J_1\ K\ J_2\ H]$, states, inputs, and input derivatives. The goal here is to keep the maximum steady state error bounded.

Definition 4. *The direction of xu_{max} – that is, the (input,state) vector that maximizes the error, is that of the singular-vector corresponding to the maximum singular-value of G, where G is the input-output transfer matrix of the error dynamics at a desired frequency [112].*

From (5.8), we introduce

$$\begin{aligned} A_e &\triangleq N^* \\ B_e &\triangleq N^*L\eta + [E^*\ J_1^*\ K^*\ J_2^*\ H^*]Q_1 - Q_2 \\ C_e &\triangleq I, \end{aligned} \tag{6.11}$$

where N^* and $[E^*\ J_1^*\ K^*\ J_2^*\ H^*]$ are a feasible solution of the estimator matrices. Define G to be the transfer matrix representation of $(A_e, B_e, C_e, 0)$ at a desired frequency. From Definition 4, xu_{max} can be obtained as the singular-vector corresponding to the maximum-singular value of G. We can then define the following Theorem.

Theorem 6. *For a system of form (2.1), if there exists an estimator of form (5.6) to make bounded-error estimations of the current and/or upcoming desired functionals of the states, then the conditions in Proposition 4 are satisfied and for the calculated* N^*, $[E^* \ {J_1}^* \ K^* \ {J_2}^* \ H^*]$, *and* xu_{max}, *a pair* (C_T, Z) *exist to satisfy the following conditions*

1.

$$Z(I - T_{1,*}{T_{1,*}}^+) - C_T{T_{1,*}}^+ + T_{2,*}{T_{1,*}}^+ - [E^* \ {J_1}^* \ K^* \ {J_2}^* \ H^*] = 0 \quad (6.12)$$

2.

$$N^*C_{1,*} \leq C_T \leq N^*C_{2,*} \quad (6.13)$$

In (6.12) and (6.13)

$$\begin{aligned} C_{1,*} &\triangleq e_1 + L\eta xu_{max} \\ C_{2,*} &\triangleq e_2 + L\eta xu_{max} \\ T_{1,*} &\triangleq Q_1 xu_{max} \\ T_{2,*} &\triangleq Q_2 xu_{max}. \end{aligned}$$

Proof. As we obtained xu_{max} to be the error-maximizing vector of a system with $[E \ J_1 \ K \ J_2 \ H] = [E^* \ {J_1}^* \ K^* \ {J_2}^* \ H^*]$ and $N = N^*$, Theorem 6 shows that with the obtained $[E \ J_1 \ K \ J_2 \ H]$ and N and with xu_{max}, still the estimation error can remain bounded (within the pre-specified values). □

Note that satisfaction of (6.12) and (6.13) are not necessary for the existence of the bounded error observer due to the fact that the designed estimator from Proposition 4 is not unique. Hence, if the designed estimator in Proposition 4 does not satisfy the conditions in Theorem 6, still several other estimators may exist which can bound the maximum error of estimation.

The sufficient condition in Theorem 6 will become a necessary and sufficient condition by applying recursion, such that the calculated xu_{max} be added to the vector XU at each stage and a new estimator be designed till all conditions in Theorem 6 are satisfied. The convergence of such a recursion can however remain an open issue. It is, therefore, easier to solve the conditions in the Proposition 4 and Theorem 6 simultaneously to obtain the necessary and sufficient condition.

Corollary 7. *Having matrix XU_r with columns randomly selected to be feasible (input,state) vectors of the process, there exists a bounded error estimator of form (5.6) to keep the estimation error within the pre-specified bounds iff the following conditions are simultaneously satisfied.*

- *There exists a random vector V, a stable N, and a matrix Z, to satisfy (6.2) and (6.3) for $XU = XU_r$.*
- *For the corresponding N^*, $[E^* \ {J_1}^* \ K^* \ {J_2}^* \ H^*]$, and the bounded xu_{max}, there exists a pair (C_T, Zn) to satisfy (6.12) and (6.13), for $Z = Z_n$.*

6.3 Application example

We consider the PKPD model described in Section 5.3, equation 5.43. The matrix F of (2.1) can be designed to make the output follow the reference trajectory.

In (5.43) and (5.44), k_{ij} and k_d are rate constants and V_1 is the volume of the plasma compartment.

The desired task which is controlling the depth of anesthesia can be defined as $Tx = \begin{bmatrix} 0 & 0 & 0 & \gamma_h (4EC_{50})^{-1} \end{bmatrix} x$, where EC_{50} is the 50% effect concentration and γ_h is the cooperativity coefficient.

In Chapter 5, we could show that by only measuring the depth of hypnosis, it would not be possible for the anesthetist to precisely predict the depth of anesthetia (DOA). This states that the SA cannot be precisely achieved by the anesthetist to perform a task on DOA while only having access to information about DOA. However, adding more information to the user interface could help to attain SA.

We modify our previous analysis in two directions:

1. In the previous chapter, we investigated the possibility that the anesthetist could estimate the current and future DOA with complete precision. However, intuitively, the anesthetist does not require to precisely reconstruct and predict the task (i.e. reconstruction and prediction of the DOA within acceptable ranges would be sufficient). Hence, we use our new technique for this analysis.

2. In the previous chapter, we assumed the anesthetist knew the precise dynamics of each patient. This, however, does not sound like a reasonable assumption. The internal estimator of the anesthetist can be considered to be formed based on a model of a nominal patient. This nominal model is the understanding of the anesthetist about an average patient in a specific category (e.g. children or adults) and is created from the real responses of various patients. In this chapter, we model the human as an estimator designed based on a nominal system, we then evaluate whether the obtained model can be used to reconstruct and predict the DOA of each patient within the desired bounds.

For the patients with average response to the administered drug, the PKPD coefficients are presented in Table 6.1. The PK parameters are estimated from [1] and the PD values are from [2]. We randomly select two sets of coefficients in Table 6.1 to form the nominal model of an average-patient that the anesthetist knows internally. We then investigate the chances that with such an internal understanding about an average patient, the anesthetist can attain SA regarding the desired task within the acceptable bounds for each patient.

Patient	k_{10}	k_{12}	k_{13}	k_{21}	k_{31}	V_1	k_d	EC_{50}	E_0	γ
1	0.0068	0.0019	0.0007	0.0009	0.0001	20.164	1.15	3.95	93.11	1.74
2	**0.0062**	**0.0019**	**0.0007**	**0.0009**	**0.0001**	**11.5058**	**1.34**	**4.24**	**92.46**	**1.90**
3	0.0062	0.0019	0.0007	0.0009	0.0001	10.0848	10.71	5.77	91.47	1.56
4	0.0061	0.0019	0.0007	0.0009	0.0001	12.331	1.12	4.84	91.6	1.55
5	0.0065	0.0019	0.0007	0.0009	0.0001	28.0782	3.84	3.97	92.91	1.62
6	0.0062	0.0019	0.0007	0.0009	0.0001	10.7266	1.89	3.57	94.58	1.57
7	**0.0062**	**0.0019**	**0.0007**	**0.0009**	**0.0001**	**10.5432**	**4.55**	**4.81**	**92.89**	**1.55**
8	0.0059	0.0019	0.0007	0.0009	0.0001	26.8164	1.46	3.71	91.68	1.75
9	0.0062	0.0019	0.0007	0.0009	0.0001	11.5975	1.16	5.44	90.30	1.52
10	0.0063	0.0019	0.0007	0.0009	0.0001	22.44	7.41	3.60	91.38	1.82

Table 6.1: Patients' parameters from [1] and [2]

For our analysis, we consider two type of measurements which include 1) only the DOA ($rank(C) = 1$ and $rank(D) = 0$), 2) the DOA, the plasma concentration; although measuring the plasma concentration is beyond current state of technology; and the automation's desired trajectory ($rank(C) = 2$ and $rank(D) = 1$). In addition, we consider three values for the prediction horizon, τ, to analyze the capability of the anesthetist to perform shorter and longer term predictions.

Our approach is to consider two of the patients as the nominal models – that is, the average-patient model that the anesthetist knows internally. We then investigate the chances that with such an internal understanding about an average patient, the anesthetist can attain SA regarding the desired task within the acceptable bounds for each patient. Note that if the conditions in Corollary 7 are all satisfied, the estimator that gives bounded delayed estimations and predictions of the desired task exists and is not necessarily unique. Hence, among all estimators that may exist based on the nominal model, some may let the anesthetist attain SA about DOA of other patients and some may not. It is, however, not clear which of the many estimators that exist is the closest model to the internal estimator of the anesthetist. We, therefore, perform a statistical analysis to determine the chances that the anesthetist can attain SA about various patients based on different internal estimators.

6.3.1 Results

For each nominal model and each combination of the measurements and τ, we design fifty estimators; if there exists any; and then evaluate whether the designed estimator is effective to attain SA about other patients.

The results are provided in Tables 6.2 and 6.3. Each table shows the chances (percentage) that the estimators designed based on a given nominal model are effective to make bounded-error estimations for other patients.

As it is expected, the estimators designed to make bounded error estimations for the nominal models 2 and 7 are effective to make correct estimations for Patients 2 and 7 (respectively) all the times.

τ (sec)	rank(C)	rank(D)	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
0.5	1	0	100	100	24	100	62	90	70	98	100	30
	2	1	5	100	2.5	52.5	5	70	25	7.5	42.5	7.5
5	1	0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	2	1	6	100	6	4	0	4	0	6	4	0
20	1	0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	2	1	0	100	0	0	0	0	0	0	0	0

Table 6.2: Percentage effectiveness of the estimator designed for nominal model P2 on estimating the task for other patients

τ (sec)	rank(C)	rank(D)	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
0.5	1	0	94	100	74	98	98	98	100	98	100	56
	2	1	0	17.5	25	17.5	0	52.5	100	0	25	0
5	1	0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	2	1	0	8	2	2	4	4	100	2	2	0
20	1	0	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	2	1	0	0	0	0	0	0	100	0	0	0

Table 6.3: Percentage effectiveness of the estimator designed for nominal model P7 on estimating the task for other patients

To discuss the results of Tables 6.2 and 6.3 in detail, we first clarify the difference between the selected values of the prediction horizon, $\tau = 0.5$, $\tau = 5$, and $\tau = 20$. The very short prediction horizon, $\tau = 0.5$, means that predicting a very short step ahead is desired and the explicit prediction of the states is not required for attaining SA. On the other hand, by $\tau = 5$ and $\tau = 20$ we mean that in order to attain situation awareness, the anesthetist is required to make explicit predictions 5 and 20 seconds in advance, respectively.

From the results obtained for $\tau = 0.5$, we can see that the estimator designed based on the nominal internal model of the anesthetist, is not necessarily capable of reconstructing and predicting the task for each individual. However, we need to notice that the possibility of making correct bounded estimations depends on the similarity between the actual and the nominal PKPD models. In addition, it can be seen that when the amount of measured information in the display is increased, it becomes less possible for the anesthetist to make correct estimations on various individuals based on the internal model. This can be due to the fact that by introducing additional measurements the internal estimator of the anesthetist is designed more specifically for the available internal nominal model.

For $\tau = 5$ and $\tau = 20$, when only the DOA is measured in the display, the conditions in Corollary 7 are not all satisfied. So, irrespective of the internal model of the anesthetist, the anesthetist cannot make correct bounded estimations of the task. In other words, it is never possible for the anesthetist to attain SA, even about the patient with the model being that of the internal nominal model. By increasing the information in the display, the internal estimators for the anesthetist can be designed to make correct estimations on the nominal model. For $\tau = 5$, in the majority of the cases, these internal estimators are not capable to let the anesthetist attain SA about other patients. When longer term prediction is required – that is, $\tau = 20$, the anesthetist can only attain SA about the DOA of the patient if s/he knows the precise model of the patient.

6.3.2 Discussion

From the results obtained in Section 6.3.1, regardless of the nominal model, the type of the measurements, or the definition of the prediction for SA, there is always a chance that the anesthetist cannot predict the task states within the desired bounds. Hence, a hazardous situation may occur at some point. The error due to the lack of SA can be as minor as putting the patient in a slightly lighter or deeper anesthesia than what is desired. It can also be very serious with the patient being put in too deep of an anesthesia. Obviously, in real world applications, where the anesthetist has access to further information about the patient through the environment, the chances that s/he cannot control and/or monitor the DOA can be much slimmer than what we have obtained here.

Due to the importance of the concept of SA in the safety of operations, and based on the results that show the existence of the cases that the anesthetist may have lack of SA about the DOA, we need to seek a way that guarantees the existence of SA for the anesthetist all the time. The solution could be providing the anesthetist with SA through a CDSS which presents predicted effect of the anesthetic drug on the patients. Two such systems are Navigator Applications Suite by GE or the SmartPilot View by Drager [113]. It is still an open issue to investigate whether with the existing uncertainties and with the differences between the PKPD models used to build these devices and the actual PKPD values of each patient, the final prediction remains in the safe bound or not.

6.4 Summary and discussion

In this chapter, the user was modeled as a bounded-error delayed functional estimator. For accomplishing a desired task safely, this estimator have to exist to reconstruct and predict a specific functional of the states of the system within pre-specified bounds.

Our method was used to investigate the important problem of safety of an anesthetized patient. Considering the anesthetist to have an internal nominal understanding about the patients, the chances that the anesthetist will be able to attain SA about the DOA of each patient during surgery was evaluated. We could show that when the available information is restricted to the displayed information, there always exists a possibility that the anesthetist cannot attain SA about the patient's DOA – that is, the understanding of the anesthetists about the depth of anesthesia of the patient is not necessarily correct. This led us to suggest incorporating automated devices which could provide the current and the predicted values of the DOA directly to the doctor.

Chapter 7

Thesis summary

7.1 Thesis contributions

The contributions in this thesis are twofold. In the body of the thesis, the effectiveness of the displayed information was investigated. In order to maintain the coherency of the techniques provided in the body, a part of the contributions, which is purely analytical, is provided in the appendices.

The body of this thesis considered the problem of evaluating the displayed information and also designing good user-interfaces for LTI systems under human or shared control. The theory of situation awareness states the importance of the comprehension and prediction of certain information before attempting to perform a task. Based on this theory, we considered the user to be a specific type of observer/predictor and evaluated the information which is required for this estimator to make correct estimations of the desired functionals of the states of the system. We introduced two main approaches for making such an evaluation of the displayed information. These approaches are applicable to systems of different orders i.e. small size systems as well as large systems with many states.

The first approach was based on subspace analysis, such that, we evaluated whether the space spanned with the combination of the states which are involved in the task can be observed and predicted by the user. For this purpose we needed to formulate two main spaces, the user-observable and the user-predictable subspaces and then investigate whether the task space is contained in these two subspaces. To determine the user-observable and the user-predictable subspaces, we considered certain limitations for the user. With such limitations, the space which is spanned by the functional of the states whose current and future values can be estimated by the user differs from the standard observable and predictable subspace.

In the second approach, we modeled the user as a specific type of observer and predictor. To create this model, we also considered certain user's specifications and limitations, including the information processing delay of the human , having access to higher derivatives of the measured states, and focusing only on the desired task. We first modeled the user as a type of observer whose goal is to make precise estimations and predictions of the task functional. We also modeled the user as an estimator whose estimation error is bounded rather than being precisely zero. Through modeling the user, we could then evaluate the existing user interfaces as well as determining the required information to be included in the display for the cases that off-line design of the displayed information is valid over the entire process.

In addition to the results on the display design, in the Appendices, we achieved some novel analytical results on observability subspaces and observer design. We evaluated the effect of higher derivatives on 1) the observability subspaces and 2) the existence of, order, and the design of functional observers.

The contributions of this thesis can be summarized as follows:

- Formulating the novel user-observable and user-predictable spaces [102, 114] (and also the delay-incorporating user-observable and the delay-incorporating user-predictable) and using them to evaluate the correctness of the displayed information [115].

- Modeling the user as a specific type of observer and predictor (for both the precise estimation [116, 117] and bounded-error estimation [118] of the task functional). Then using this model to evaluate the existing display or to design a new display.

- Determining the effect of the availability of higher derivatives of the input and the output on the observability subspace [119] and on the existence and order of the functional observer[120].

7.2 Possible future directions

Since the presented research topic in this thesis is very novel and not much work has been done on the subject, there are many directions that can be taken to make current results stronger and more suitable for real world applications.

1. **Incorporating nonlinearities as well as uncertainties and noise:**

   In this thesis we have evaluated the existing display content and have suggested a design technique for designing the displays for LTI systems with no noise and uncertainty. This assumption of having a non-noisy and deterministic LTI system is, however, too simplistic when it comes to real world applications.

   First, throughout the thesis we considered the possible availability of the derivatives of input and output signals under the assumption of having non-noisy input and output signals. However, this assumption have to be relaxed at some point to give us more realistic results. For the case of having noisy signals, the techniques have to be extensively modified since it is not straight forward to deal with the derivatives of the noise.

   Besides, consider the PKPD model of the patient we discussed in Sections 5.3 and 6.3. Not only this model is affected by external noise terms such as measurement noise, the system matrices also are not precise and have a degree of uncertainty associated with them. This may affect the obtained results to a great deal such that while we believe that the estimation error remains bounded during the operation, it may grow beyond the acceptable bounds.

   In addition to the uncertainties and noises, considering the system to be a LTI is rather simplistic. On one hand, the linearized systems are not the most realistic presentation of each system. On the other hand, in addition to all different sources of noise, the linearization itself introduces more uncertainties to the model.

Hence, robust and non-linear analysis of the user as an observer/predictor is required to provide us with a guaranteed or at least a more realistic bound of error of estimation. The main benefit of modeling the user as a specific type of robust observer is that this observer can be designed to be robust to the effect of bounded parametric uncertainties as well as the noise. This is, in fact, how a trained and experienced user would behave.

2. **Relaxing some assumptions:**

   Several assumptions were made in the thesis to help with obtaining the initial model of the user.

   For instance, so far, we considered a specific category of systems in which the user-interface was the only source of information for the user. Incorporating the effect of information from the environment on the required information content of the user-interface is an important work which is necessary in order to make our framework more applicable to realistic cases.

   In addition we made some fundamental assumptions including the limited role of mental model on simply understanding the dynamics and also non-existence of mind wandering. We believe that at further stages of this project, the mentioned simplifications have to be relaxed to provide us with a more comprehensive framework.

3. **Coming up with a rigorous technique to determine the required information to be displayed:**

   In Chapter 5 we suggested a heuristic algorithm to determine the display content. This algorithm, however, is only valid if the operating conditions and task are non-changing during the entire process. If the operating conditions and/or the desired task change at some point, a real time estimation of the displayed information is necessary. For such a real time determination of information, we require an analytical formulation of an effective display that can be easily updated during the process. In addition, a numerical solution with fast convergence may be helpful to determine the displayed information in real-time.

As has been discussed in Section 5.2.4, a solution to the optimization problem (5.42) gives us an analytical tool to determine the information content of the display. Other techniques may also exist and might be more suitable to determine the required information, either analytically or numerically.

4. **Model validation:**

   Having developed a model for the user attaining SA, it is also required to come up with a technique to validate such a model. The model that we suggested for the human is not unique. In addition, it is obviously not feasible to validate each and every model that could be created. However, it can be a good idea to investigate whether the display is actually rejected when our technique claims that it would be. For instance, a basic experiment could be to ask the users to accomplish a simple task with a rejected display.

   Overall, designing a well-developed test for validating the suggested human model is not only important but it is also very challenging.

Bibliography

[1] A. Absalom and G. Kenny, "'Paedfusor'pharmacokinetic data set," *British Journal of Anaesthesia*, vol. 95, no. 1, pp. 110–110, 2005.

[2] K. van Heusden, J. M. Ansermino, K. Soltesz, S. Khosravi, N. West, and G. A. Dumont, "Quantification of the variability in response to propofol administration in children," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 9, pp. 2521–2529, 2013.

[3] M. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors*, vol. 37, pp. 32–64, 1995.

[4] M. Heymann and A. Degani, "On abstractions and simplifications in the design of human-automation interfaces," NASA Ames Research Center, Moffett Field, CA, NASA Technical Memorandum 211397, 2002.

[5] P. Ladkin and H. Sogame, "Aircraft accident investigation report 96-5," http://sunnyday.mit.edu/accidents/nag-contents.html, July 1996.

[6] N. Matni and M. Oishi, "Reachability-based abstraction for an aircraft landing under shared control," in *Proceedings of the American Control Conference*, Seattle, WA, June 2008, pp. 2278–2284.

[7] N. Sarter, D. Woods, and C. Billings, "Automation surprises," in *Handbook of Human Factors and Ergonomics.* NY: John Wiley and Sons, Inc., 1999, pp. 1295–1327.

[8] N. Leveson, L. Pinnel, S. Sandys, S. Koga, and J. Reese, "Analyzing software specifications for mode confusion potential," in *Proceedings of*

a Workshop on Human Error and System Development, C. Johnson, Ed. Glasgow, Scotland: Glasgow Accident Analysis Group, March 1997, pp. 132–146.

[9] J. Crow, D. Javaux, and J. Rushby, "Models and mechanized methods that integrate human factors into automation design," in *International Conference on Human-Computer Interaction in Aeronautics*, Toulouse, France, September 2000.

[10] R. Butler, S. Miller, J. Potts, and V. Carreno, "A formal methods approach to the analysis of mode confusion," in *Proceedings of the AIAA/IEEE Digital Avionics Systems Conference*, 1998, pp. C41/1–C41/8.

[11] M. Oishi, I. Mitchell, A. M. Bayen, and C. J. Tomlin, "Invariance-preserving abstractions of hybrid systems: Application to user interface design," *IEEE Transactions on Control System Technology*, vol. 16, no. 2, pp. 229–244, March 2008.

[12] M. Paull and S. Unger, "Minimizing the number of states in incompletely specified sequential switching functions," *IRE Transactions on Electronic Computers*, pp. 356–367, September 1959.

[13] C. Lesire and C. Tessier, "Estimation and conflict detection in human controlled systems," in *Hybrid Systems: Computation and Control*, ser. LNCS 3927, J. Hespanha and A. Tiwari, Eds. Springer Verlag, March 2006, pp. 407–420.

[14] N. Moray and T. B. Sheridan, "On sont les neiges d'antan?" *Human Performance, Situation Awareness, and Automation: Current Research and Trends HPSAA II, Volumes I and II*, p. 1, 2013.

[15] M. Endsley, "Situation models: An avenue to the modeling of mental models," in *14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society*, Santa Monica, CA, USA, 2000.

[16] M. R. Endsley, *Designing for Situation Awareness: An Approach to User-Centered Design.* Taylor & Francis US, 2003.

[17] J. R. Busemeyer and A. Diederich, *Cognitive Modeling.* Sage, 2010.

[18] T. A. Polk and C. M. Seifert, *Cognitive Modeling.* MIT Press, 2002.

[19] W. D. Gray and E. M. Altmann, "Cognitive modeling and human-computer interaction," *Karwowski [341]*, pp. 387–391, 2001.

[20] F. I. Craik and R. S. Lockhart, "Levels of processing: A framework for memory research," *Journal of verbal learning and verbal behavior*, vol. 11, no. 6, pp. 671–684, 1972.

[21] R. M. Shiffrin and M. Steyvers, "A model for recognition memory: Rem? retrieving effectively from memory," *Psychonomic Bulletin & Review*, vol. 4, no. 2, pp. 145–166, 1997.

[22] N. Chater and M. Oaksford, *Rational Models of Cognition.* Wiley Online Library, 2003.

[23] N. Chater, J. B. Tenenbaum, and A. Yuille, "Probabilistic models of cognition: Conceptual foundations," *Trends in Cognitive Sciences*, vol. 10, no. 7, pp. 287–291, 2006.

[24] T. L. Griffiths, C. Kemp, and J. B. Tenenbaum, "Bayesian models of cognition," 2008.

[25] D. McRuer and D. Weir, "Theory of manual vehicular control," *Ergonometrics*, vol. 12, no. 4, pp. 599–633, 1969.

[26] D. McRuer, R. Allen, D. Weir, and R. Klein, "New results in driver steering control model," *Human Factors*, vol. 19, no. 4, pp. 381–397, 1977.

[27] R. Hess and A. Modjtahedzadeh, "A control theoretic model of driver steering behavior," *IEEE Control Systems Magazine*, vol. 10, no. 5, pp. 3–8, 1990.

[28] R. Hess, "Model for human use of motion cues in vehicular control," *Journal of Guidance, Control, and Dynamics*, vol. 13, no. 3, pp. 476–482, 1990.

[29] L. Chen and A. Galip Ulsoy, "Identification of a driver steering model and model uncertainty, from driving simulator data," *Journal of Dynamic Systems, Measurement, and Control*, vol. 123, no. 4, pp. 623–629, 2001.

[30] M. Gabibulayev and B. Ravani, "A stochastic form of a human driver steering dynamics model," *Journal of Dynamic Systems, Measurement, and Control*, vol. 129, no. 3, pp. 322–336, 2007.

[31] D. M. Gaba, S. K. Howard, and S. D. Small, "Situation awareness in anesthesiology," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 37, no. 1, pp. 20–31, 1995.

[32] J. J. Salerno, M. L. Hinman, and D. M. Boulware, "A situation awareness model applied to multiple domains," in *Defense and Security*, 2005, pp. 65–74.

[33] F. E. Ritter, J. L. Bittner, S. E. Kase, R. Evertsz, M. Pedrotti, and P. Busetta, "Cojack: A high-level cognitive architecture with demonstrations of moderators, variability, and implications for situation awareness," *Biologically Inspired Cognitive Architectures*, vol. 1, pp. 2–13, 2012.

[34] D. Kleinman, S. Baron, and W. Levison, "An optimal control model of human response part i: Theory and validation," *Automatica*, vol. 6, no. 3, pp. 357–369, 1970.

[35] ——, "A control theoretic approach to manned-vehicle systems analysis," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 824–832, 1971.

[36] E. G. Gai and R. E. Curry, "A model of the human observer in failure

detection tasks," *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-6, no. 2, pp. 85 –94, feb. 1976.

[37] P. Wewerinke, "Model of the human observer and decision maker: Theory and validation," *Automatica*, vol. 19, no. 6, pp. 693–696, 1983.

[38] D. Luenberger, "Observing the state of a linear system," *IEEE Transactions on Military Electronics*, vol. 8, no. 2, pp. 74–80, 1964.

[39] ——, "An introduction to observers," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 596–602, 1971.

[40] G. Basile and G. Marro, "On the observability of linear time-invariant systems with unknown inputs," *Journal of Optimization Theory and Applications*, vol. 3, no. 6, pp. 410–415, 1969.

[41] R. Kalman, "On the general theory of control systems," *IRE Transactions on Automatic Control*, vol. 4, no. 3, p. 110, 1959.

[42] P. Murdoch, "Observer design for a linear functional of the state vector," *IEEE Transactions on Automatic Control*, vol. 18, no. 3, pp. 308 – 310, jun 1973.

[43] J. Moore and G. Ledwich, "Minimal order observers for estimating linear functions of a state vector," *IEEE Transactions on Automatic Control*, vol. 20, no. 5, pp. 623 – 632, oct 1975.

[44] C.-C. Tsui, "A new algorithm for the design of multifunctional observers," *IEEE Transactions on Automatic Control*, vol. 30, no. 1, pp. 89 – 93, jan 1985.

[45] S.-Y. ZHANG, "Functional observer and state feedback," *International Journal of Control*, vol. 46, no. 4, pp. 1295–1305, 1987.

[46] C.-C. Tsui, "On the order reduction of linear function observers," *IEEE Transactions on Automatic Control*, vol. 31, no. 5, pp. 447 – 449, may 1986.

[47] M. Aldeen and H. Trinh, "Reduced-order linear functional observer for linear systems," *IEE Proceedings on Control Theory and Applications*, vol. 146, no. 5, pp. 399 –405, sep 1999.

[48] F. Rotella and I. Zambettakis, "Minimal single linear functional observers for linear systems," *Automatica*, vol. 47, no. 1, pp. 164–169, 2011.

[49] M. Darouach, "Existence and design of functional observers for linear systems," *IEEE Transactions on Automatic Control*, vol. 45, no. 5, pp. 940 –943, may 2000.

[50] H. Trinh and T. Fernando, "On the existence and design of functional observers for linear systems," in *International Conference on Mechatronics and Automation*. IEEE, 2007, pp. 1974–1979.

[51] T. Fernando, H. M. Trinh, and L. Jennings, "Functional observability and the design of minimum order linear functional observers," *IEEE Transactions on Automatic Control*, vol. 55, no. 5, pp. 1268 –1273, may 2010.

[52] M. Darouach, "On the functional observers for linear descriptor systems," *Systems & Control Letters*, vol. 61, no. 3, pp. 427–434, 2012.

[53] L. Jennings, T. Fernando, and H. Trinh, "Existence conditions for functional observability from an eigenspace perspective," *IEEE Transactions on Automatic Control*, vol. 56, no. 12, pp. 2957–2961, 2011.

[54] M. Hou, A. Pugh, and P. Muller, "Disturbance decoupled functional observers," *IEEE Transactions on Automatic Control,*, vol. 44, no. 2, pp. 382–386, 1999.

[55] H. Trinh and Q. Ha, "Design of linear functional observers for linear systems with unknown inputs," *International Journal of Systems Science*, vol. 31, no. 6, pp. 741–749, 2000.

[56] H. Trinh, T. Fernando, and S. Nahavandi, "Design of reduced-order functional observers for linear systems with unknown inputs," *Asian Journal of Control*, vol. 6, no. 4, pp. 514–520, 2004.

[57] H. Deng and H.-X. Li, "Functional observers for linear systems with unknown inputs," *Asian Journal of Control*, vol. 6, no. 4, pp. 462–468, 2004.

[58] M. Darouach, "Functional observers for systems with unknown inputs," in *Proceedings of the 16th International Symposium on Mathematical Theory of Networks and Systems*, Leuven,Belgium, July 2004.

[59] T. Fernando, S. MacDougall, V. Sreeram, and H. Trinh, "Existence conditions for unknown input functional observers," *International journal of control*, vol. 86, no. 1, pp. 22–28, 2013.

[60] M. Darouach and M. Boutayeb, "Design of observers for descriptor systems," *IEEE Transactions on Automatic Control*, vol. 40, no. 7, pp. 1323 –1327, jul 1995.

[61] H. Trinh, "Linear functional state observer for time-delay systems," *International Journal of Control*, vol. 72, no. 18, pp. 1642–1658, 1999.

[62] M. Darouach, "Linear functional observers for systems with delays in state variables," *IEEE Transactions on Automatic Control*, vol. 46, no. 3, pp. 491–496, 2001.

[63] J.-e. Feng, J. Lam, and S. Xu, "Finite-time functional observers for descriptor systems," *International Journal of Control, Automation and Systems*, vol. 7, no. 3, pp. 341–347, 2009.

[64] M. Darouach, "Unknown inputs observers design for delay systems," *Asian Journal of Control*, vol. 9, no. 4, pp. 426–434, 2007.

[65] W. Leong, H. Trinh, and T. Fernando, "A practical functional observer scheme for interconnected time-delay systems," *International Journal of Control*, no. ahead-of-print, pp. 1–11, 2015.

[66] H. Xu, Z. Lin, and A. Makur, "The existence and design of functional observers for two-dimensional systems," *Systems & Control Letters*, vol. 61, no. 2, pp. 362–368, 2012.

[67] K. Emami, T. Fernando, B. Nener, H. Trinh, and Y. Zhang, "A functional observer based fault detection technique for dynamical systems," *Journal of the Franklin Institute*, vol. 352, no. 5, pp. 2113–2128, 2015.

[68] D. Huong, H. Trinh, H. Tran, and T. Fernando, "Approach to fault detection of time-delay systems using functional observers," *Electronics Letters*, vol. 50, no. 16, pp. 1132–1134, 2014.

[69] J. Moreno, "Simultaneous observation of linear systems: A state-space interpretation," *IEEE Transactions on Automatic Control*, vol. 50, no. 7, pp. 1021–1025, 2005.

[70] H. Trinh, P. Teh, and Q. Ha, "A common disturbance decoupled observer for linear systems with unknown inputs," *International Journal of Automation and Control*, vol. 2, no. 2, pp. 286–297, 2008.

[71] J. Jin, M.-J. Tahk, and C. Park, "Time-delayed state and unknown input observation," *International Journal of Control*, vol. 66, no. 5, pp. 733–746, 1997.

[72] A. Saberi, A. Stoorvogel, and P. Sannuti, "Exact, almost and optimal input decoupled (delayed) observers," *International Journal of Control*, vol. 73, no. 7, pp. 552–581, 2000.

[73] S. Sundaram and C. Hadjicostis, "Delayed observers for linear systems with unknown inputs," *IEEE Transactions on Automatic Control*, vol. 52, no. 2, pp. 334–339, 2007.

[74] H. Trinh and T. Fernando, *Functional Observers for Dynamical Systems.* Springer, 2012.

[75] C. E. Billings, *Aviation Automation: The Search for a Human-Centered Approach*, 1997.

[76] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A model for types and levels of human interaction with automation," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 30, no. 3, pp. 286–297, 2000.

[77] L. Sherry, M. Feary, P. Polson, and E. Palmer, "What's it doing now?: Taking the covers off autopilot behavior," in *11th International Symposium on Aviation Psychology*, Columbus, OH, USA, 2001.

[78] N. Cowan, "What are the differences between long-term, short-term, and working memory?" *Progress in Brain Research*, vol. 169, pp. 323–338, 2008.

[79] A. Degani and H. Heymann, "Formal verification of human-automation interaction," *Human Factors*, vol. 44, no. 1, pp. 28–43, 2002.

[80] D. G. Mitchell and D. H. Klyde, "Recommended practices for exposing pilot-induced oscillations or tendencies in the development process," *USAF Developmental Test and Evaluation Summit*, vol. 6810, 2004.

[81] G. Jamieson and K. Vicente, "Designing effective human-automation-plant interfaces: A control-theoretic perspective," *Human Factors*, vol. 47, no. 1, pp. 12–34, 2005.

[82] T. Sheridan, *Telerobotics, Automation, and Supervisory Control.* Cambridge, MA: MIT Press, 1992.

[83] M. Cummings and K. Thornburg, "Paying attention to the man behind the curtain," *IEEE Pervasive Computing*, vol. 10, pp. 58–62, 2011.

[84] C. C. Macadam, "Understanding and modeling the human driver," *Vehicle System Dynamics*, vol. 40, no. 1-3, pp. 101–134, 2003.

[85] M. R. Endsley, "Theoretical underpinnings of situation awareness: A critical review," *Situation Awareness Analysis and Measurement*, pp. 3–32, 2000.

[86] ——, *Designing for Situation Awareness: An Approach to User-Centered Design.* CRC Press, 2012.

[87] E. Awh, E. Vogel, and S.-H. Oh, "Interactions between attention and working memory," *Neuroscience*, vol. 139, no. 1, pp. 201–208, 2006.

[88] D. Fougnie, "The relationship between attention and working memory," *New Research on Short-Term Memory*, pp. 1–45, 2008.

[89] S. J. Luck, G. F. Woodman, and E. K. Vogel, "Event-related potential studies of attention," *Trends in Cognitive Sciences*, vol. 4, no. 11, pp. 432–440, 2000.

[90] M. C. Pasolunghi, C. Cornoldi, and S. De Liberto, "Working memory and intrusions of irrelevant information in a group of specific poor problem solvers," *Memory & Cognition*, vol. 27, no. 5, pp. 779–790, 1999.

[91] B. Carretti, C. Cornoldi, R. De Beni, and M. Romano, "Updating in working memory: A comparison of good and poor comprehenders," *Journal of Experimental Child Psychology*, vol. 91, no. 1, pp. 45–66, 2005.

[92] B. Carretti, C. Cornoldi, R. De Beni, and P. Palladino, "What happens to information to be suppressed in working-memory tasks? short and long term effects," *The Quarterly Journal of Experimental Psychology Section A*, vol. 57, no. 6, pp. 1059–1084, 2004.

[93] R. De Beni and P. Palladino, "Decline in working memory updating through ageing: Intrusion error analyses," *Memory*, vol. 12, no. 1, pp. 75–89, 2004.

[94] J. Smallwood and J. W. Schooler, "The restless mind." *Psychological Bulletin*, vol. 132, no. 6, p. 946, 2006.

[95] J. M. Smallwood, S. F. Baracaia, M. Lowe, and M. Obonsawin, "Task unrelated thought whilst encoding information," *Consciousness and Cognition*, vol. 12, no. 3, pp. 452–484, 2003.

[96] J. Roskam, *Airplane Flight Dynamics and Automatic Flight Controls*. DARcorporation, 2001.

[97] E. L. Duke, R. F. Antoniewicz, and K. D. Krambeer, *Derivation and Definition of a Linear Aircraft Motion*, NASA, 1988, reference Publication 1207.

[98] "Flight crew operating manual, A310," Air India.

[99] J. Fax and R. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1465 – 1476, 2004.

[100] L. Lafferriere, A. Williams, J. Caughman, and J. Veerman, "Decentralized control of vehicle formations," *Systems and and Control Letters*, vol. 54, pp. 899–910, 2005.

[101] X. Ding, M. Powers, M. Egerstedt, S. Young, and T. Balch, "Executive decision support," *Robotics Automation Magazine, IEEE*, vol. 16, no. 2, pp. 73 –81, 2009.

[102] N. Eskandari and M. Oishi, "Computing observable and predictable subspaces to evaluate user-interfaces of lti systems under shared control," in *Proceedings of the IEEE conference on Systems, Man and Cybernetics*, Alaska, USA, October 2011, pp. 2803–2808.

[103] W. L. Brogan, "Modem control theory," *Prentice Hall, Englewood*, 1991.

[104] C. D. Wickens, "Temporal limits of human information processing: A developmental study." *Psychological Bulletin*, vol. 81, no. 11, p. 739, 1974.

[105] G. R. Loftus, T. A. Busey, and J. W. Senders, "Providing a sensory basis for models of visual information acquisition," *Perception & Psychophysics*, vol. 54, no. 4, pp. 535–554, 1993.

[106] R. Hosman and H. Stassen, “Pilot’s perception in the control of aircraft motions,” *Control engineering practice*, vol. 7, no. 11, pp. 1421–1428, 1999.

[107] E. Fioratou, R. Flin, R. Glavin, and R. Patey, “Beyond monitoring: Distributed situation awareness in anaesthesia,” *British Journal of Anaesthesia*, vol. 105, no. 1, pp. 83–90, 2010.

[108] C. M. Schulz, M. R. Endsley, E. F. Kochs, A. W. Gelb, and K. J. Wagner, “Situation awareness in anesthesia: Concept and research,” *Anesthesiology*, vol. 118, no. 3, pp. 729–742, 2013.

[109] J.-O. Hahn, G. A. Dumont, and J. M. Ansermino, “Observer-based strategies for anesthesia drug concentration estimation.”

[110] ——, “Closed-loop anesthetic drug concentration estimation using clinical-effect feedback,” *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 1, pp. 3–6, 2011.

[111] J. Schüttler and H. Ihmsen, “Population pharmacokinetics of propofol: a multicenter study,” *Anesthesiology*, vol. 92, no. 3, pp. 727–738, 2000.

[112] S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control: Analysis and Design*. Wiley New York, 2007, vol. 2.

[113] T. Gin, “Clinical pharmacology on display,” *Anesthesia & Analgesia*, vol. 111, no. 2, pp. 256–258, 2010.

[114] R. Cortez, D. Tolic, I. Palunko, N. Eskandari, R. Fierro, M. Oishi, and J. Wood, “A hybrid framework for user-guided prioritized search and adaptive tracking of maneuvering targets via cooperative uavs,” *Progress in Astronautics and Aeronautics*, 2012.

[115] N. Eskandari, G. A. Dumont, and Z. J. Wang, “Delay-incorporating observability and predictability analysis of safety-critical continuous-time systems,” *IET Control Theory and Applications*, vol. 9, pp. 1692–1699(7), July 2015.

[116] N. Eskandari, Z. J. Wang, and G. A. Dumont, “On the existence and design of functional observers for lti systems, with application to user modeling,” *Asian Journal of Control*, 2014.

[117] N. Eskandari, G. A. Dumont, and Z. J. Wang, “An observer/predictor based model of the user for attaining situation awareness,” *IEEE Transactions on Human-Machine Systems*, 2015.

[118] ——, “Bounded-error delayed functional observer/predictor: Existence and design,” *IEEE Transactions on Biomedical Engineering, submitted*.

[119] N. Eskandari, M. M. Oishi, and Z. J. Wang, “Observability analysis of continuous-time lti systems with limited derivative data,” *Asian Journal of Control*, vol. 16, no. 2, pp. 623–627, 2014.

[120] N. Eskandari, Z. J. Wang, and G. Dumont, “Modeling the user as an observer to determine display information requirements,” in *In Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*. IEEE, 2013, pp. 267–272.

[121] J. Cieslak, D. Henry, and A. Zolghadri, “Fault tolerant flight control: From theory to piloted flight simulator experiments,” *Control Theory Applications, IET*, vol. 4, no. 8, pp. 1451 –1461, august 2010.

Appendix A

On the effect of higher derivatives on the observability subspace

Our main goal in this appendix is to show the effect of providing information about the input, output, and a set of their derivatives on state reconstruction. We, therefore, formulate the observability subspace with up to the i^{th} input derivatives available, and also up to the j^{th} output derivatives available.

In Section A.2, we come up with a formula to analyze the unknown input observability subspace. We then evaluate how the availability of the derivatives of input and output signals may affect the results in Section A.3.

A.1 Problem formulation

Consider a continuous-time LTI system of the form

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bw(t) \\ y(t) &= Cx(t), \end{aligned} \tag{A.1}$$

with state vector $x(t) \in \mathbb{R}^n$, continuous-time output $y(t) \in \mathbb{R}^p$, and coefficient matrices A, B, and C with compatible dimensions. In Section A.2 we consider $w(t) \in \mathbb{R}^m$ to be a continuous-time unknown input and in Section A.3, $w(t)$ is a known input with derivatives which can be unknown. For simplicity, in the rest of the chapter, we omit the index (t) from $x(t)$, $y(t)$, $w(t)$, and their derivatives.

In [40] the authors show that the observability subspace of (2.1) (with unknown input) is the least $(A^T, \mathcal{F}^{\perp})$-conditioned invariant containing $\mathcal{R}(C^T)$ where $\mathcal{F} \triangleq \mathcal{N}(B^T)^{\perp}$ (i.e. the column space of B). In this chapter we evaluate the effect of lacking information about some derivatives of input and output, rather than unknown input, on the observability subspace of the system.

A.2 Unknown-input observability subspace

In a slightly different approach from [40], we re-formulate the unknown-input observability subspace of (2.1) and obtain an equation which is easier to solve and extend. To do so, we introduce the orthogonal projection, $P_1 \in \mathbb{R}^{p \times p}$, onto the left null space of CB. With pre-multiplying P_1 in the first derivative of the output equation, we obtain

$$\begin{aligned} P_1 \dot{y} &= P_1(CAx + CBw) \\ &= P_1 CAx, \end{aligned} \tag{A.2}$$

which leads to the set of states which can be reconstructed regardless of the values of unknown inputs. Clearly, when CB is of full column rank, P_1 is a zero matrix.

Continuing in a similar fashion with higher derivatives, and introducing further projection matrices, P_k $\forall k \in \{2, \cdots, n-1\}$, to remove the input from up to the k^{th} derivative of the output equation (i.e. P_k is an orthogonal projection onto the space $\mathcal{N}(B^T(A^T)^{k-1}C^T \prod_{i=1}^{k-1} P_k{}^T)$), we can have

$$\prod_{k=i}^{1} P_k y^{(i)} = \prod_{k=i}^{1} P_k C A^i x, \tag{A.3}$$

where $\prod_{k=i}^{1} P_k \triangleq P_i P_{i-1} \cdots P_1$. Putting the output equation and equation

(A.3) (for $i \leq (n-1)$) together, we have

$$\begin{bmatrix} y \\ P_1\dot{y} \\ P_2P_1\ddot{y} \\ \vdots \\ \prod_{k=n-1}^{1} P_k y^{(n-1)} \end{bmatrix} = O_p x \tag{A.4}$$

with

$$O_p \triangleq \begin{bmatrix} C \\ P_1CA \\ P_2P_1CA^2 \\ \vdots \\ \prod_{k=n-1}^{1} P_k CA^{n-1} \end{bmatrix}. \tag{A.5}$$

Theorem 7. *The unknown-input observability subspace of (2.1), which itself is a subset of the observability space of the mentioned system, is a time-invariant space*

$$\mathcal{O}_{\mathrm{UI}} \triangleq \mathcal{R}(C^T) \oplus \sum_{i=1}^{n-1} \mathcal{R}\left((A^T)^i C^T \prod_{k=1}^{i} P_k{}^T\right), \tag{A.6}$$

where

$$P_k = p(\mathcal{N}(B^T(A^T)^{k-1}C^T \prod_{i=1}^{k-1} P_i{}^T))$$

and $p(\mathcal{M})$ means the projection onto the space $\mathcal{M}$.

Proof. By definition, $\mathcal{O}_{\mathrm{UI}}$ is the largest subspace in which states that span it can be reconstructed without the knowledge of the input and its derivatives. From (A.4), the states which span the row space of O_p can be reconstructed, irrespective of the unknown inputs. Therefore, the unknown-input observability subspace is the row space of O_p,

$$\begin{aligned}\mathcal{R}\left(O_p^T\right) &= \mathcal{R}\left(C^T\right) \oplus \mathcal{R}\left(A^T C^T {P_1}^T\right) \oplus \mathcal{R}\left((A^T)^2 C^T {P_1}^T {P_2}^T\right) \oplus \cdots \oplus \\ &\quad \mathcal{R}\left((A^T)^{n-1} C^T \prod_{k=1}^{n-1} {P_k}^T\right) \\ &= \mathcal{R}(C^T) \oplus \sum_{i=1}^{n-1} \mathcal{R}\left((A^T)^i C^T \prod_{k=1}^{i} {P_k}^T\right).\end{aligned} \tag{A.7}$$

Hence, (A.6) is obtained. □

Since, for any two matrices A and B with compatible dimensions, we have $\mathcal{R}(AB) = A\mathcal{R}(B)$ and $\mathcal{R}(AB) \subseteq \mathcal{R}(A)$, and also from $\prod_{k=i}^{1} P_k C A^{i-1} B = 0$, we can obtain

$$\mathcal{R}\left((A^T)^i C^T \prod_{k=1}^{i} {P_k}^T\right) \subseteq \left[A^T(\mathcal{R}\left((A^T)^{i-1} C^T\right) \cap \mathcal{N}(B^T))\right]. \tag{A.8}$$

From (A.7) and (A.8), we have

$$\begin{aligned}\mathcal{R}({O_p}^T) \subseteq\; & \mathcal{R}(C^T) \oplus A^T(\mathcal{R}(C^T) \cap \mathcal{N}(B^T)) \oplus \\ & A^T\left(\mathcal{R}(A^T C^T) \cap \mathcal{N}(B^T)\right) \oplus \cdots \oplus \\ & A^T(\mathcal{R}\left((A^T)^{n-2} C^T\right) \cap \mathcal{N}(B^T)),\end{aligned} \tag{A.9}$$

therefore,

$$\begin{aligned}\mathcal{O}_{\mathrm{UI}} &\subseteq \mathcal{R}(C^T) \oplus A^T \sum_{i=0}^{n-2} \left(\mathcal{R}((A^T)^i C^T) \cap \mathcal{N}(B^T)\right) \\ &\subseteq \mathcal{R}(C^T) \oplus A^T(\mathcal{O} \cap \mathcal{N}(B^T)),\end{aligned} \tag{A.10}$$

where $\mathcal{O}$ represents the observability subspace of (2.1).

As a result of (A.10), we can provide the following remark.

Remark . For an unknown-input observable system for which $\mathcal{O}_{\mathrm{UI}} = \mathbb{R}^n$,

$$\mathcal{R}(C^T) \oplus A^T(\mathcal{O} \cap \mathcal{N}(B^T)) = \mathbb{R}^n. \tag{A.11}$$

□

A.3 Observability subspace with limited information about the input and output derivatives

In this section we obtain an equation that determines how providing more information about input and output derivatives, affects the respective observability subspace of the system. For this purpose $w(t)$ is considered to be known but with some unknown derivatives.

We first obtain the relationship between two observability subspaces, 1) $\mathcal{O}_{i-1}$: The observability subspace with up to the $(i-1)^{th}$ input derivatives available and 2) $\mathcal{O}_i$: The observability subspace with up to the i^{th} input derivatives available.

Assume that we have information about input and up to its $(i-1)^{th}$ derivative. So, for the $(i+1)^{th}$ derivative of the output equation and the projection matrix P_1, we can show that

$$P_1 y^{(i+1)} = P_1 C A^{i+1} x + P_1 C A^i B w + \cdots + P_1 C A B w^{(i-1)}. \tag{A.12}$$

Proceeding with higher derivatives, we pre-multiply each equation with a projection matrix (as in the proof of Theorem 7), then

$$Y = O_{pd} x + HW, \tag{A.13}$$

where H is a lower-triangular Toeplitz matrices of compatible size and

$$O_{pd} \triangleq \begin{bmatrix} C & CA & \cdots & P_1CA^{i+1} & \cdots & \prod_{k=n-1-i}^{1} P_kCA^{n-1} \end{bmatrix}^T. \quad \text{(A.14)}$$

In (A.13),

$$\begin{aligned} Y &\triangleq \begin{bmatrix} y & \dot{y} & \cdots & P_1y^{(i+1)} & \cdots & \prod_{k=n-1-i}^{1} P_ky^{(n-1)} \end{bmatrix}^T \\ W &\triangleq \begin{bmatrix} w & \dot{w} & \cdots & w^{(i-1)} \end{bmatrix}^T. \end{aligned} \quad \text{(A.15)}$$

The states which span the row-space of O_{pd} can be reconstructed irrespective of unknown values of input derivatives. Hence, with $\mathcal{O}_{i-1} = \mathcal{R}({O_{pd}}^T)$,

$$\mathcal{O}_{i-1} = \mathcal{R}(C^T) \oplus \sum_{j=1}^{i} \mathcal{R}\left((A^T)^j C^T\right) \oplus \sum_{j=i+1}^{n-1} \mathcal{R}\left((A^T)^j C^T \prod_{k=1}^{n-i} {P_k}^T\right). \quad \text{(A.16)}$$

Theorem 8. *For (2.1), the observability subspace with up to the i^{th} input derivatives available is*

$$\begin{aligned} \mathcal{O}_i &= \mathcal{R}(C^T) \oplus A^T \mathcal{O}_{i-1} \\ \mathcal{O}_0 &\triangleq \mathcal{R}(C^T) \oplus A^T \mathcal{O}_{\mathrm{UI}}, \end{aligned} \quad \text{(A.17)}$$

where $\mathcal{O}_0$ is the observability subspace with only 0^{th} input derivative available.

Proof. Pre-multiplying P_1 in the $(i+2)^{th}$ derivative of the output, eliminates the unknown derivative of input. Taking more derivatives and introducing more projection matrices to remove the unknown values and putting the results together, we obtain (A.13) with

$$O_{pd} \triangleq \begin{bmatrix} C \\ CA \\ \vdots \\ P_1 C A^{i+2} \\ \vdots \\ \prod_{k=n-2-i}^{1} P_k C A^{n-1} \end{bmatrix}. \tag{A.18}$$

Therefore

$$\begin{aligned} \mathcal{O}_i &= \mathcal{R}(C^T) \oplus \sum_{j=1}^{i+1} \mathcal{R}\left((A^T)^j C^T\right) \oplus \sum_{j=i+2}^{n-1} \mathcal{R}\left((A^T)^j C^T \prod_{k=1}^{n-i} {P_k}^T\right) \\ &= \mathcal{R}(C^T) \oplus A^T(\mathcal{R}(C^T) \oplus \\ &\quad \sum_{j=1}^{i} \mathcal{R}\left((A^T)^j C^T\right) \oplus \sum_{j=i+1}^{n-1} \mathcal{R}\left((A^T)^j C^T \prod_{k=1}^{n-i} {P_k}^T\right)) \\ &= \mathcal{R}(C^T) \oplus A^T \mathcal{O}_{i-1} \end{aligned} \tag{A.19}$$

and note that the 0^{th} derivative case reduces to the result in Theorem 7. Hence,

$$\mathcal{O}_0 \triangleq \mathcal{R}(C^T) \oplus A^T \mathcal{O}_{\mathrm{UI}}. \tag{A.20}$$

□

In Figure A.1 we show how providing information about input and its derivatives can affect the observability subspace of the system. Note that the two dimensional representation of subspaces is just a simplification. As has been mathematically shown in Theorem 8, by providing information about the input and its derivatives, a larger part of the state space can be reconstructed. Note that the largest subspace that can be reconstructed is the observability subspace, $\mathcal{O}$, for which all information about input, output,

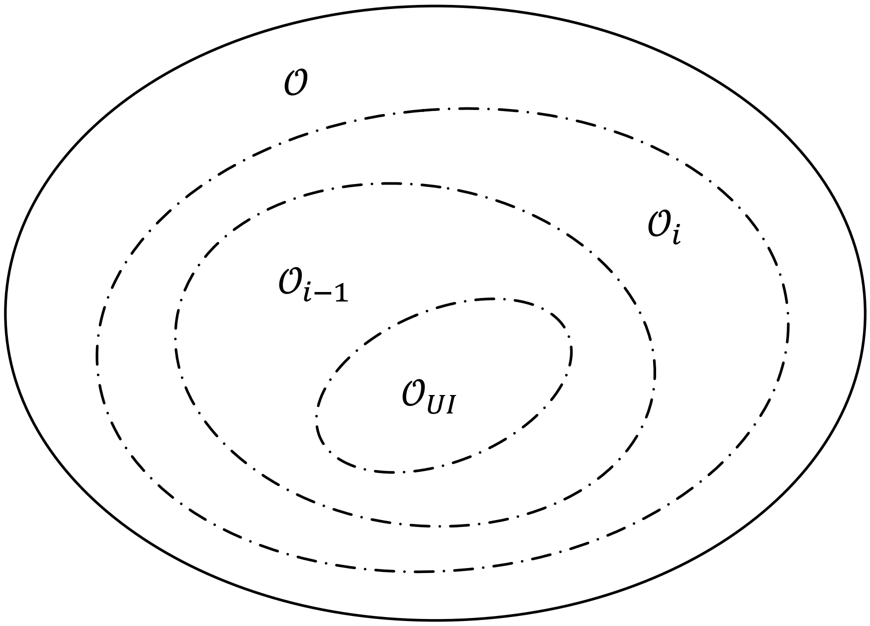


Figure A.1: The effect of providing information about input and its derivatives on the observability subspace. We could show that adding information about input and its derivatives can result in larger observability space. The dashed-dotted lines represent containment.

and their derivatives up to the $(n-1)^{th}$ derivatives is available.

If in addition to limited information about higher input derivatives, the information about output derivatives is also limited, a smaller part of the state space can be reconstructed.

Corollary 8. *For (2.1), the observability subspace with up to the i^{th} input derivatives and up to the j^{th} output derivatives available, (with $j > i$), is*

$$\begin{aligned} \mathcal{O}_{i,j} &= \mathcal{R}(C^T) \oplus A^T \mathcal{O}_{i-1,j} \\ \mathcal{O}_{0,j} &\triangleq \mathcal{R}(C^T) \oplus A^T \mathcal{O}_{\mathcal{UI},j}, \end{aligned} \tag{A.21}$$

with

$$\mathcal{O}_{\mathcal{UI},j} = \mathcal{R}(C^T) \oplus A^T \sum_{i=0}^{j-1} \left(\mathcal{R}((A^T)^i C^T) \cap \mathcal{N}(B^T) \right).$$

Proof. Having limited information about output derivatives will affect (A.6) by changing the upper margin of the summation. On the other hand, the recursive part of equation (A.17) which shows the effect of providing further information about input derivatives will not be affected. Hence, (A.21) is obtained.

Since the i^{th} derivative of input will not show up until taking $(i+1)$ derivatives from the output equation, it is necessary to have $j > i$. Otherwise, every i in (A.21) should be replaced by $(j-1)$.

□

Remark. It is straight forward to show that Theorems 7 and 8 and also Corollary 8 will remain the same for a system model (2.1) with both an unknown input (or an input with unknown derivatives) and a known input whose derivatives are entirely known. □

A.4 Example

Consider the linearized longitudinal dynamics of a Boeing 747 in trimmed level flight [121], with state $x = [q, V, \alpha, \theta, h]$ consisting of pitch rate q, airspeed V, angle of attack α, pitch angle θ, and altitude h. We consider $u(t) = 0$ and assume that the user applies the input $w(t) = \delta_{ih}(t)$ which represents the deflection of the horizontal tail. System matrices are provided in (A.24).

Consider the case in which the measurement is limited to information about the pitch angle, therefore, $C = [0\ 0\ 0\ 1\ 0]$. The observable subspace of this system spans $\mathbb{R}^5$, hence, all states are observable. From Theorem 7, the unknown-input observable subspace (A.6) is

$$\mathcal{O}_{\text{UI}} = \text{span}\left(e_1, e_4\right). \tag{A.22}$$

We now evaluate how having information about the input and its first two derivatives but not having information about the third and forth derivatives of input can affect the observable subspace. So, from (A.17)

$$\begin{aligned} \mathcal{O}_0 &= \operatorname{span}(e_1, e_3, e_4) \\ \mathcal{O}_1 &= \operatorname{span}(e_1, e_3, e_2 - 0.0086e_5, e_4) \end{aligned} \tag{A.23}$$

and $\mathcal{O}_2 = \mathbb{R}^5$ spans the entire state-space.

Now consider a case that the information about the derivatives of output is also limited. From (A.21), $\mathcal{O}_{\mathcal{UI},0} = \mathcal{O}_{\mathrm{UI}}$. Since from Corollary 8 it is required to have $j > i$, the new observable subspace will span $\mathbb{R}^5$ for $j \geq 3$.

Essentially, when measuring the pitch angle of an aircraft with dynamics provided in (A.24), the first two derivatives of input and the first three derivatives of output are enough to reconstruct the entire state space and higher derivatives of these signals do not provide additional information about the states of the system.

$$
\begin{aligned}
A &= \begin{bmatrix}
-6.6926\times10^{-1} & -8.6\times10^{-6} & -8.856\times10^{-1} & 0 & -3.45\times10^{-6} \\
-1.6179\times10^{-1} & -7.588\times10^{-3} & 4.9965 & -9.8 & 4.59\times10^{-5} \\
1.0084 & -1.0036\times10^{-3} & -6.735\times10^{-1} & 0 & 5.9\times10^{-6} \\
1 & 0 & 0 & 0 & 0 \\
0 & 0 & -1.338\times10^{2} & 1.338\times10^{2} & 0
\end{bmatrix}, \\
B &= \begin{bmatrix} -4.5944\times10^{-2} & 0 & -1.912\times10^{-3} & 0 & 0 \end{bmatrix}^T
\end{aligned}
\tag{A.24}
$$

Appendix B

On the effect of higher derivatives on the existence of and design of functional observers

In this appendix we focus on investigating the effects of availability of the extended input and output signals (i.e., higher derivatives of input and output signals in this chapter) on the existence of a generic functional observer for LTI systems.

In Section B.2, we derive the existence conditions of a generic functional observer of form (B.6) for system (B.1). We provide an estimate of the required order of the functional observer in B.3, and we suggest a design procedure for such an observer in Section B.3.1.

B.1 Problem formulation

The results provided in this chapter are a specific case of those in Chapter 5. Here, we focus specifically on an extension of the model used in [58],

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + \alpha F d(t) \\ y(t) &= Cx(t) + G_1 u(t) + \alpha G_2 d(t), \end{aligned} \tag{B.1}$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^{m_u}$ is the known input, $d(t) \in \mathbb{R}^{m_d}$ is the unknown input or the disturbance, $y(t) \in \mathbb{R}^p$ is the output vector and the matrices A, B, F, C, G_1, and G_2 have compatible dimensions. In

(B.1), $\alpha \in \{0, 1\}$, where $\alpha = 0$ for having no unknown inputs and $\alpha = 1$ for having unknown inputs. Instead of introducing the parameter α, F can be alternatively considered as being a zero matrix for systems without unknown inputs.

Assuming that we have information about the inputs and the outputs as well as their derivatives, our goal is to evaluate whether or not there exists a functional observer to reconstruct the linear functional

$$z_0(t) = L_0 x(t), \quad z_0(t) \in \mathbb{R}^r, \tag{B.2}$$

where $L_0 \in \mathbb{R}^{r\times n}$.

To reduce the required order of the functional observer, we introduce the design parameters γ and λ. For the system with γ derivatives of outputs available, we redefine the set of outputs as $Y_{0:\gamma} = [y^T(t), \dot{y}^T(t), \cdots, {y^{(\gamma)}}^T(t)]^T$, thus

$$Y_{0:\gamma}(t) = O_\gamma x(t) + M_{1,0:\gamma} U_{0:\gamma}(t) + \alpha M_{2,0:\gamma} D_{0:\gamma}(t) \tag{B.3}$$

where O_γ is the observability matrix and $M_{1,0:\gamma}$ and $M_{2,0:\gamma}$ are Toeplitz matrices as follows,

$$\begin{aligned}
O_\gamma &= [C^T, A^T C^T, \cdots, A^{\gamma T} C^T]^T, \\
M_{1,0:\gamma} &= \begin{bmatrix} G_1 & 0 & \cdots & 0 \\ CB & G_1 & \cdots & 0 \\ CAB & CB & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{\gamma-1}B & CA^{\gamma-2}B & \cdots & G_1 \end{bmatrix}, \\
M_{2,0:\gamma} &= \begin{bmatrix} G_2 & 0 & \cdots & 0 \\ CF & G_2 & \cdots & 0 \\ CAF & CF & \cdots & G_2 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{\gamma-1}F & CA^{\gamma-2}F & \cdots & G_2 \end{bmatrix}, \\
U_{0:\gamma}(t) &= [u^T(t), \dot{u}^T(t), \cdots, {u^{(\gamma)}}^T(t)]^T, \\
D_{0:\gamma}(t) &= [d^T(t), \dot{d}^T(t), \cdots, {d^{(\gamma)}}^T(t)]^T,
\end{aligned} \tag{B.4}$$

Clearly, for all $i \in \mathbb{N}$, we have $M_{1,i:\gamma+i} = M_{1,0:\gamma}$ and $M_{2,i:\gamma+i} = M_{2,0:\gamma}$. We also can write these matrices in recursive form as

$$\begin{aligned} M_{1,0:\gamma+1} &= \left[\begin{array}{c|c} G_1 & 0 \\ \hline O_\gamma B & M_{1,0:\gamma} \end{array}\right], \\ M_{2,0:\gamma+1} &= \left[\begin{array}{c|c} G_2 & 0 \\ \hline O_\gamma F & M_{2,0:\gamma} \end{array}\right], \end{aligned} \tag{B.5}$$

and

$$\begin{aligned} U_{0:\gamma+1}(t) &= [U^T{}_{0:\gamma}(t), {u^{(\gamma+1)}}^T(t)]^T, \\ D_{0:\gamma+1}(t) &= [D^T{}_{0:\gamma}(t), {d^{(\gamma+1)}}^T(t)]^T, \end{aligned}$$

Based on the extended outputs, $Y_{0:\gamma}$, and the extended inputs, $U_{0:\lambda} = [u^T(t), \dot{u}^T(t), \cdots, {u^{(\lambda)}}^T(t)]^T$, our goal is to design a functional observer to reconstruct the desired functional, $z_0(t)$. Therefore the observer dynamics is

$$\begin{aligned} \dot{w}(t) &= Nw(t) + JY_{0:\gamma} + HU_{0:\lambda}, \\ \hat{z}(t) &= w(t) + EY_{0:\gamma}, \end{aligned} \tag{B.6}$$

where $w(t) \in \mathbb{R}^{r+\mathcal{X}}$. Note that, designing a stable observer is equivalent to determining matrices J, H, E, and a stable N with compatible dimensions.

Similar to what we did in Chapter 5, we also assume that in order to reconstruct the functional $z_0(t) = L_0 x(t)$, it is also necessary to reconstruct the functional $Rx(t)$ where $R \in \mathbb{R}^{\mathcal{X}\times n}$. Hence, we introduce the extended functional as

$$z(t) = \begin{bmatrix} L_0 \\ R \end{bmatrix} x(t), \tag{B.7}$$

such that $L^T = [{L_0}^T, R^T]^T$ is of full row rank. For cases that $L_0 x(t)$ can be reconstructed directly, we have $R = \emptyset$.

Choosing $Q \triangleq EO_\gamma - L$ and $K = J - NE$, we have the estimation error

$$\begin{aligned} e(t) &= \hat{z}(t) - z(t) \\ &= w(t) + EY_{0:\gamma}(t) - Lx(t) \\ &= w(t) + Qx(t) + EM_{1,0:\gamma}U_{0:\gamma} + \alpha EM_{2,0:\gamma}D_{0:\gamma} \end{aligned} \tag{B.8}$$

with the following dynamics

$$\begin{aligned} \dot{e}(t) &= Ne(t) + (QA + JO_\gamma - NQ)x(t) + \\ &\quad KM_{1,0:\gamma}U_{0:\gamma}(t) + EM_{1,1:\gamma+1}U_{1:\gamma+1} + \\ &\quad HU_{0:\lambda} + QBu(t) + \alpha KM_{2,0:\gamma}D_{0:\gamma} + \\ &\quad EM_{2,1:\gamma+1}D_{1:\gamma+1} + \alpha QFd(t). \end{aligned} \tag{B.9}$$

From (B.5), (B.9) can be written as

$$\begin{aligned} \dot{e}(t) &= Ne(t) + (QA + JO_\gamma - NQ)x(t) + \\ &\quad (K \begin{bmatrix} G_1 \\ O_{\gamma-1}B \end{bmatrix} + QB + H_a)u(t) + \\ &\quad (KM_a + EM_{1,0:\gamma})U_{1:\gamma+1} + H_bU_{1:\lambda} + \\ &\quad \alpha(K \begin{bmatrix} G_2 \\ O_{\gamma-1}F \end{bmatrix} + QF)d(t) + \\ &\quad \alpha(KM_b + EM_{2,0:\gamma})D_{1:\gamma+1}, \end{aligned} \tag{B.10}$$

where

$$\begin{aligned} M_a &\triangleq \begin{bmatrix} 0 & 0_{p\times m_u} \\ M_{1,0:\gamma-1} & 0 \end{bmatrix}, \\ M_b &\triangleq \begin{bmatrix} 0 & 0_{p\times m_d} \\ M_{2,0:\gamma-1} & 0 \end{bmatrix}, \end{aligned} \tag{B.11}$$

and

$$\begin{aligned} H_a &\triangleq H \qquad \lambda = 0, \\ \begin{bmatrix} H_a & H_b \end{bmatrix} &\triangleq H \qquad \lambda > 0. \end{aligned} \tag{B.12}$$

By choosing $\beta = \max(\lambda, \gamma + 1)$, we can write

$$\begin{aligned} \left[\begin{array}{c|c} M_{1,1} & M_{1,2} \end{array} \right] &\triangleq M_{1,0:\gamma}, \\ \left[\begin{array}{c|c} M_{2,1} & M_{2,2} \end{array} \right] &\triangleq M_{a}, \end{aligned} \tag{B.13}$$

where $M_{1,2} \in \mathbb{R}^{(\gamma\times p)\times((\beta-\lambda)\times m_u)}$ and $M_{2,2} \in \mathbb{R}^{(\gamma\times p)\times((\beta-\lambda)\times m_u)}$. Note that for $\gamma \leq \lambda$, we have $M_{1,1} \triangleq M_{1,0:\gamma}$ and $M_{2,1} \triangleq M_a$. Also, for $\lambda = 0$, $M_{1,2} \triangleq M_{1,0:\gamma}$ and $M_{2,2} \triangleq M_a$.

To have an asymptotically stable observer, from (B.9) and (B.13), the matrix N has to be stable and we need to satisfy

$$\begin{aligned} EO_\gamma A + KO_\gamma &= LA - NL, \\ \alpha(EO_\gamma F + K \begin{bmatrix} G_2 \\ O_{\gamma-1}F \end{bmatrix}) &= \alpha LF, \\ \alpha(EM_{2,0:\gamma} + KM_b) &= 0, \end{aligned} \tag{B.14}$$

and

$$\begin{aligned} EO_\gamma B + K \begin{bmatrix} G_1 \\ O_{\gamma-1}B \end{bmatrix} + H_a &= LB, \\ EM_{1,1} + KM_{2,1} + H_b &= 0, \qquad \lambda > 0. \end{aligned} \tag{B.15}$$

In addition to (B.14) and (B.15), for $\gamma > \lambda$, we also need to satisfy

$$EM_{1,2} + KM_{2,2} = 0. \tag{B.16}$$

In summary, to design a stable functional observer of form (B.6) for the reconstruction of the functional $z_0(t)$ in system (B.1), we do the following two steps:

1. We evaluate the existence of such a functional observer by checking whether there exists a stable matrix N to satisfy (B.14), (B.15), and (B.16).

2. After estimating the required order of the observer, we design the functional observer by determining the observer matrices to satisfy

(B.14), (B.15), and (B.16) for an stable N.

It is worth mentioning that the process of taking derivatives from the input and the output signals can amplify the high frequency components which is undesirable. To reduce the adverse effect of these noisy components of the extended signals, low-pass filtering might be needed. As our focus here is on non-noisy systems and non-noisy derivatives of the signals, determining an appropriate filtering technique is out of the scope of the current chapter (e.g., in our main application of human-automation interaction, we assume that the user is capable of evaluating the rate of change of the available signals).

B.2 Existence conditions of a functional observer

For LTI systems with known and/or unknown inputs and with available derivatives of inputs and outputs, we obtain the necessary and sufficient conditions for the existence of a functional observer.

Recall that we have considered having a full row rank functional $Lx(t)$, with $Lx(t) = \begin{bmatrix} L_0 \\ R \end{bmatrix} x(t)$, whose components are $L_0x(t)$ and $Rx(t)$. Therefore reconstructing $Lx(t)$ results in the reconstruction of $L_0x(t)$.

With a similar approach to that used in [58], we can write (B.14) and (B.16) as

$$\begin{bmatrix} E & K \end{bmatrix} T_1 = T_2, \tag{B.17}$$

where

$$T_1 = \begin{bmatrix} O_\gamma A E_1 & \alpha O_\gamma F & \alpha M_{2,0:\gamma} & M_{1,2} \\ O_\gamma E_1 & \alpha \begin{bmatrix} G_2 \\ O_{\gamma-1} F \end{bmatrix} & \alpha M_b & M_{2,2} \end{bmatrix}$$

and

$$T_2 = \begin{bmatrix} LAE_1 & \alpha LF & 0 & 0 \end{bmatrix},$$

with E_1 being selected so that $LE_1 = 0$.

Considering (B.17), this equation has a solution if and only if $\text{span}({T_2}^T) \subseteq \mathcal{R}({T_1}^T)$ or equivalently

$$rank \begin{bmatrix} T_2 \\ T_1 \end{bmatrix} = rank[T_1]. \tag{B.18}$$

Theorem 9. *An observer with order $(r + \mathcal{X})$ exists to reconstruct the functional $z_0 = L_0 x(t)$ of the system (B.1) (with γ and λ being the design parameters) iff there exists a matrix $R \in \mathbb{R}^{\mathcal{X} \times n}$ for which*

-
$$rank(LHS_1) = rank(RHS) \tag{B.19}$$

- *For all $s \in \mathbb{C}$,*
$$rank(LHS_2) = rank(RHS) \tag{B.20}$$

where

$$LHS_1 \triangleq \begin{bmatrix} L_0 A & \alpha L_0 F & 0 & 0 \\ RA & \alpha RF & 0 & 0 \\ O_\gamma A & \alpha O_\gamma F & \alpha M_{2,0:\gamma} & M_{1,2} \\ O_\gamma & \alpha \begin{bmatrix} G_2 \\ O_{\gamma-1} F \end{bmatrix} & \alpha M_b & M_{2,2} \\ L_0 & 0 & 0 & \\ R & 0 & 0 & 0 \end{bmatrix}, \tag{B.21}$$

$$LHS_2 \triangleq \begin{bmatrix} sL_0 - L_0 A & -\alpha L0F & 0 & 0 \\ sR - RA & -\alpha RF & 0 & 0 \\ O_\gamma A & \alpha O_\gamma F & \alpha M_{2,0:\gamma} & M_{1,2} \\ O_\gamma & \alpha \begin{bmatrix} G_2 \\ O_{\gamma-1} F \end{bmatrix} & \alpha M_b & M_{2,2} \end{bmatrix}, \tag{B.22}$$

$$RHS \triangleq \begin{bmatrix} O_\gamma A & \alpha O_\gamma F & \alpha M_{2,0:\gamma} & M_{1,2} \\ O_\gamma & \alpha \begin{bmatrix} G_2 \\ O_{\gamma-1}F \end{bmatrix} & \alpha M_b & M_{2,2} \\ L_0 & 0 & 0 & 0 \\ R & 0 & 0 & 0 \end{bmatrix}. \tag{B.23}$$

Proof.

- Proof of (B.19).

  By selecting H_1 to satisfy $LH_1 = I$ and post multiplying the full row rank matrix

  $$S_1 = \begin{bmatrix} H_1 & E_1 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix}$$

  in (B.21) and (B.23), we can show that satisfaction of (B.19) for a matrix R of rank $\mathcal{X}$ is equivalent to satisfaction of (B.18) and guarantees existence of solution for (B.14) and (B.16).

- Proof of (B.20).

  Assume that condition (B.19) is satisfied, then

  $$\begin{bmatrix} E & K \end{bmatrix} = T_2 {T_1}^+ + Z(I - T_1 {T_1}^+) \tag{B.24}$$

  are the solutions of (B.17) for arbitrary matrix Z with compatible dimensions. Note that when $(I - T_1 {T_1}^+) = 0$, matrices E and K are not affected by Z and can be uniquely obtained as $\begin{bmatrix} E & K \end{bmatrix} = T_2 {T_1}^+$.

  It is also required to satisfy the stability of the observer (B.6). We can rewrite the first equation in (B.14) as follows:

  $$N = LAH_1 - \begin{bmatrix} E & K \end{bmatrix} \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} H_1. \tag{B.25}$$

From (B.24) and (B.25), we can have $N = \Lambda - Z\Gamma$ where

$$\begin{aligned} \Lambda &= LAH_1 - T_2{T_1}^{+} \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} H_1, \\ \Gamma &= (I - T_1{T_1}^{+}) \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} H_1. \end{aligned} \tag{B.26}$$

Clearly the eigenvalues of N can be placed at any desired values iff $rank \begin{bmatrix} sI - \Lambda \\ \Gamma \end{bmatrix} = r + \mathcal{X},\ \forall s \in \mathbb{C}.$

By post-multiplying S_1 in (B.22), we have

$$\begin{bmatrix} sI - LAH_1 & -T_2 \\ \begin{bmatrix} O_\gamma AH_1 \\ O_\gamma H_1 \end{bmatrix} & T_1 \end{bmatrix}. \tag{B.27}$$

Now choose a S_2 with full column rank

$$S_2 = \begin{bmatrix} I & T_2{T_1}^{+} \\ 0 & (I - T_1{T_1}^{+}) \\ 0 & T_1{T_1}^{+} \end{bmatrix} \tag{B.28}$$

and a S_3 with full row rank

$$S_3 = \begin{bmatrix} I & 0 \\ -{T_1}^{+} \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} H_1 & I \end{bmatrix}. \tag{B.29}$$

By pre-multiplying S_2 and post-multiplying S_3 in (B.27), the rank does not change and we can finally obtain $rank(LHS_2) \triangleq \Psi_1$ to be

$$rank \begin{bmatrix} sL_0 - L_0A & -\alpha L_0 F & 0 & 0 \\ sR - RA & -\alpha RF & 0 & 0 \\ O_\gamma A & \alpha O_\gamma F & \alpha M_{2,0:\gamma} & M_{1,2} \\ O_\gamma & \alpha \begin{bmatrix} G_2 \\ O_{\gamma-1}F \end{bmatrix} & \alpha M_b & M_{2,2} \end{bmatrix}, \tag{B.30}$$

where

$$\Psi_1 \triangleq rank \begin{bmatrix} sI - \Lambda \\ \Gamma \end{bmatrix} + rank(T_1). \tag{B.31}$$

Similar to the proof of (B.19), if we post-multiply S_1 in (B.23), we can show that its rank is

$$rank \begin{bmatrix} O_\gamma A & \alpha O_\gamma F & \alpha M_{2,0:\gamma} & M_{1,2} \\ O_\gamma & \alpha \begin{bmatrix} G_2 \\ O_{\gamma-1}F \end{bmatrix} & \alpha M_b & M_{2,2} \\ L_0 & 0 & 0 & 0 \\ R & 0 & 0 & 0 \end{bmatrix} = \Psi_2, \tag{B.32}$$

where

$$\Psi_2 \triangleq r + \mathcal{X} + rank(T_1). \tag{B.33}$$

From (B.30) and (B.32) we have $rank \begin{bmatrix} sI - \Lambda \\ \Gamma \end{bmatrix} = r + \mathcal{X}$ iff (B.20) is satisfied.

□

We can show that the results of Theorems in [49] and [58] can be easily obtained from Theorem 9 by having $\alpha = 0$ and $\alpha = 1$ respectively, with $\gamma = 0$ and $\lambda = 0$.

Theorem 9 requires finding a matrix R with a low rank (with unknown number of rows) which satisfies (B.19) and (B.20) simultaneously. The matrix R of lowest rank can be determined through an iterative algorithm, as presented in Algorithm 1.

Algorithm 1 Determining matrix R

1: **if** (B.19) and (B.20) are satisfied with $R = \bar{\mathbf{0}}$ **then**
2: $\quad R = \bar{\mathbf{0}}$
3: **else**
4: $\quad$ Define

$$\mathcal{R} = \{\{e_i\} | i \in \{1, ..., n\} \quad \& \quad \text{span}(e_i) \nsubseteq \text{span}({L_0}^T)\}$$

and $j = 1$.
5: $\quad R_j = \{v | \text{span}(v) \subseteq \text{span}(\mathcal{R}) \ \& \ rank(v) = j\}$.
6: $\quad$ Obtain R such that

$$(L_0, R) = \{(L_0, w) | w \in R_j \ \& \ (L_0, w) \ satisfy \ (B.19)\&(B.20)\}.$$

7: $\quad$ **if** $R = \emptyset$ and $j < n$ **then**
8: $\quad\quad j = j + 1$
9: $\quad\quad$ go to 5.
10: $\quad$ **end if**
11: **end if**

As computing R for systems of high order can be numerically hard, we suggest a sufficient (but not necessary) condition in Proposition 5 below for the existence of a functional observer for system (B.1) which does not depend on the unknown matrix R.

To obtain Proposition 5, we first explain a few notations and terms. We have $\left[\begin{array}{c|c} Y_1 & Y_2 \end{array}\right]$, with Y_1 having n columns, to be the transpose of the basis of $\mathcal{N}\left(\alpha \begin{bmatrix} F \\ O_\gamma F \end{bmatrix}^T\right)$.

We calculate Y_N and Y_M such that the rows of $[Y_N | Y_M]$ are selected from the rows of $[Y_1|Y_2]$ and

$$\begin{aligned} \text{span}({Y_N}^T) &\subseteq \mathcal{N}\left(\begin{bmatrix} L_0 \\ PO_\gamma \end{bmatrix}\right), \\ \text{span}({Y_M}^T) &\subseteq \mathcal{N}(\begin{bmatrix} \alpha M_{2,0:\gamma} & M_{1,2} \end{bmatrix}^T), \end{aligned} \tag{B.34}$$

where $P = P_2 P_1$. Matrices P_1 and P_2 are projection matrices onto the left null space of $M_{1,0:\gamma}$ and $\alpha P_1 M_{2,0:\gamma}$ respectively. Note that $PY_{0:\gamma}(t) = PO_\gamma x(t)$ is the combination of outputs with no dependency on the inputs.

Also we choose M such that

$$\text{span}(M^T) \subseteq \mathcal{N}(\begin{bmatrix} \alpha M_{2,0:\gamma} & M_{1,2} \end{bmatrix}^T)$$

and to satisfy

$$\alpha(L_0 F + MO_\gamma F) = 0. \tag{B.35}$$

We choose $R = KY_N$, where K is an identity matrix with compatible dimension. Equation (B.34) shows that rows of R are selected to be linearly independent from the functional $L_0 x(t)$ and from PO_γ – that is, the available states which have no dependency on inputs. From (B.35), the rank of (B.21) can be written as

$$rank(LHS_1) = rank \left[\begin{array}{c|c} \begin{matrix} L_0 A + MO_\gamma A \\ KY_N A + KY_M O_\gamma A \end{matrix} & 0 \\ \hline \begin{matrix} O_\gamma A \\ O_\gamma \end{matrix} & T \\ \hline \begin{matrix} L_0 \\ KY_N \end{matrix} & 0 \end{array}\right], \tag{B.36}$$

where

$$T = \begin{bmatrix} \alpha O_\gamma F & \alpha M_{2,0:\gamma} & M_{1,2} \\ \alpha \begin{bmatrix} G_2 \\ O_{\gamma-1} F \end{bmatrix} & \alpha M_b & M_{2,2} \end{bmatrix}. \tag{B.37}$$

We choose $v = [v_1 \ v_2]^T$ to be a full row rank matrix such that $v_1 T$ has full row rank equal to the rank of T and $v_2 T = 0$ (e.g. $v_1 = T^+$ where T^+ is the pseudo-inverse of matrix T and $\text{span}(v_2) = \mathcal{N}(T^T)$). Hence, from (B.36),

$$\begin{aligned} rank(LHS_1) &= rank \left[\begin{array}{c|c} v_1 \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} & v_1 T \\ \hline v_2 \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} & \\ L_0 A + M O_\gamma A & \\ K Y_N A + K Y_M O_\gamma A & 0 \\ L_0 & \\ K Y_N & \end{array} \right] \\ &= q_1 + rank(T), \end{aligned} \tag{B.38}$$

where

$$q_1 = rank \begin{bmatrix} v_2 \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} \\ L_0 A + M O_\gamma A \\ K Y_N A + K Y_M O_\gamma A \\ L_0 \\ K Y_N \end{bmatrix}. \tag{B.39}$$

Having similar v_1 and v_2, for (B.23) we have

$$rank(RHS) = rank \left[\begin{array}{c|c} v_1 \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} & v_1 T \\ \hline v_2 \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} & \\ L_0 & 0 \\ KY_N & \end{array} \right] = q_2 + rank(T), \qquad \text{(B.40)}$$

where

$$q_2 = rank \begin{bmatrix} v_2 \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} \\ L_0 \\ KY_N \end{bmatrix}. \qquad \text{(B.41)}$$

Proposition 5. *For system (B.1), a functional observer of form (B.6) exists to reconstruct the functional $L_0 x(t)$ if*

- *Condition 1:*

$$rank \left[\begin{array}{c} L_0 A + M O_\gamma A \\ Y_N A + Y_M O_\gamma A \\ \hline v_2 \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} \\ L_0 \\ Y_N \end{array} \right] = rank \begin{bmatrix} v_2 \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} \\ L_0 \\ Y_N \end{bmatrix}, \qquad \text{(B.42)}$$

- *Condition 2: For all* $s \in \mathbb{C}$,

$$rank \left[\begin{array}{c} v_2 \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} \\ \hline L_0(sI - A) - MO_\gamma A \\ Y_N(sI - A) - Y_M O_\gamma A \end{array} \right] = rank \begin{bmatrix} v_2 \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} \\ L_0 \\ Y_N \end{bmatrix}. \quad \text{(B.43)}$$

with v_1, v_2, P, M, Y_N, *and* Y_M *already defined.*

Proof. • Proof of (B.42).

Earlier we have proved (B.19). We have also shown that for $R = KY_N$, equation (B.38) is an equivalent to the rank of (B.21) and equation (B.40) is equivalent to the rank of (B.23). Hence, for the above R, (B.19) and the condition $q_1 = q_2$ are equivalent.

Through proof with contrapositive, we now assume that there exists no K such that (B.19) can be satisfied when $R = KY_N$, therefore, $q_1 \neq q_2|_{\forall K}$. However, we can clearly see that for $K = I$, $q_1 = q_2$. Thus, there exists a matrix R, not exclusively $R = Y_N$, that satisfies (B.19) if (B.42) is satisfied.

- Proof of (B.43).

The rank of (B.22) can be written as

$$rank \left[\begin{array}{c|c} v_1 \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} & v_1 T \\ \hline v_2 \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} & \\ L_0(sI - A) - MO_\gamma A & 0 \\ KY_N(sI - A) - KY_M O_\gamma A & \end{array} \right], \quad \text{(B.44)}$$

hence,

$$\begin{aligned} rank(LHS_2) &= q_3 + rank(T), \\ & \quad \forall s \in \mathcal{C}, \end{aligned} \tag{B.45}$$

where

$$q_3 = rank \begin{bmatrix} v_2 \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} \\ L_0(sI - A) - MO_\gamma A \\ KY_N(sI - A) - KY_M O_\gamma A \end{bmatrix}. \tag{B.46}$$

We have already obtained (B.23) to be equal to (B.40). Hence, from (B.40) and (B.45), finding a matrix R to satisfy (B.20) is equivalent to finding a K for which ($q_3 = q_2$) for all ($s \in \mathcal{C}$). Satisfaction of (B.43) is essentially equivalent to having $(q_3 = q_2)|_{K=I}$ for all ($s \in \mathcal{C}$).

From both parts of the proof, we can show that (B.42) and (B.43) are sufficient for the existence of a functional observer (B.6) to reconstruct $L_0x(t)$ in system (B.1). □

Remark. Note that, if Y_N is a zero matrix, there exists no linearly independent combinations of the states of the system to be added to L_0. Therefore, the existence problem of a state reconstructor is reduced to existence of a reconstructor with the same order as the desired functional $z_0(t)$. □

B.3 Estimating the order of the functional observer

It is clear that the rank of (B.23) is always smaller than or equal to the rank of (B.21), i.e., $q_2 \leq q_1$ for all K. To estimate the required order of the unknown-input functional observer, we first need to determine a matrix K_a such that, for $K = K_a$, the rank of (B.23) is equal to $q_1|_{K=K_a} + rank(T)$, where $q_1|_{K=K_a} + rank(T)$ is the rank of (B.21) for $K = K_a$. Equivalently, from (B.41) and the desired condition $q_1 = q_2$, we need to find a matrix K_a such that

$$rank\begin{bmatrix} v_2\begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix} \\ L_0 \\ K_a Y_N \end{bmatrix} = q_1|_{K=K_a}. \tag{B.47}$$

In order to simplify (B.47), we choose K_a such that

$$Y_N{}^T\mathcal{R}(K_a{}^T) \subseteq \mathcal{N}(v_2\begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix}). \tag{B.48}$$

Thus, we have

$$\begin{aligned} \mathcal{X}_a &\triangleq rank(K_a Y_N) \\ &= q_1|_{K=K_a} + rank(T) - rank(v_2\begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix}). \end{aligned} \tag{B.49}$$

As $q_1 \leq q$ for $q \triangleq (q_1|_{K=I})$, to satisfy $q_1 = q_2$, we need to have $K_a \in \mathbb{R}^{\mathcal{X}_a \times n_{YN}}$, where n_{YN} is the number of rows of Y_N, such that

$$\mathcal{X}_a \leq q + rank(T) - rank(v_2\begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix}). \tag{B.50}$$

Equation (B.50) provides an upper bound for the number of rows to be added to the desired functional so that the extended functional satisfies (B.19).

On the other hand, if we find a matrix K_a for which $R = K_a Y_N$ satisfies $q_1 = q_2$, we still need to satisfy $q_2 = q_3$. For this purpose, we need to have a matrix $K_b \in \mathbb{R}^{\mathcal{X}_b \times n}$ where

$$\begin{aligned} \mathcal{X}_b &\triangleq q_2|_{K=K_{a,b}} + rank(T) - \min_{s\in eig\Lambda}(rank(C_p)) \\ &\leq q + rank(T) - \min_{s\in eig\Lambda}(rank(C_p)), \end{aligned} \tag{B.51}$$

where

$$K_{a,b} = \begin{bmatrix} K_a \\ K_b \end{bmatrix} \tag{B.52}$$

and

$$C_p = \begin{bmatrix} sL_0 - L_0A & \alpha L0F & 0 & 0 \\ sK_aY_N - K_aY_NA & \alpha K_aY_NF & 0 & 0 \\ O_\gamma A & \alpha O_\gamma F & \alpha M_{2,0:\gamma} & M_{1,2} \\ O_\gamma & \alpha \begin{bmatrix} G_2 \\ O_{\gamma-1}F \end{bmatrix} & \alpha M_b & M_{2,2} \end{bmatrix}. \tag{B.53}$$

To estimate the minimum order of the unknown-input functional observer, required for the reconstruction of $L_0x(t)$, we suggest the following steps:

1. From (B.50), obtain $\mathcal{X}_a$ and define $\mathcal{L}_1$ to be a set of all feasible $(L_1 \triangleq K_aY_N) \in \mathbb{R}^{(r+\mathcal{X}_a)\times n}$ whose corresponding K_a's satisfy $q_1 = q_2$.

2. Find

$$M_m = \max_{L_1\in\mathcal{L}_1}\left(\min_{s\in eig\Lambda}(rank(C_p))\right) \tag{B.54}$$

   and denote the corresponding optimal solution as L_m.

3. The required order of the observer is therefore

$$\mathcal{X} = \mathcal{X}_a + \mathcal{X}_m, \tag{B.55}$$

   with $\mathcal{X}_m \leq q + rank(T) - M_m$.

4. The functional to be reconstructed is

$$L = \begin{bmatrix} L_m \\ K_mY_N \end{bmatrix}, \tag{B.56}$$

with $K_m \in \mathbb{R}^{\mathcal{X}_m \times rows\ of\ Y_N}$ being selected so that $K = \begin{bmatrix} K_a \\ K_m \end{bmatrix}$ satisfies $q_1 = q_3$.

We can now introduce Proposition 6 as an alternative of Theorem 9 in subspace form.

Proposition 6. *An observer with order $(r + \mathcal{X})$ exists to reconstruct $L_0 x(t)$ in (B.1) (with γ and λ being the design parameters) iff*

- *a matrix K_a exists to satisfy*

$$ {r_1}^T \mathcal{R}({K_1}^T) \cap \mathcal{N}(P_c) = {r_L}^T \mathcal{R}({K_1}^T) \cap \mathcal{N}(P_c) \tag{B.57} $$

where

$$ K_1 = \begin{bmatrix} I & 0 \\ 0 & K_a \end{bmatrix}, \quad K_a \in \mathbb{R}^{\mathcal{X}_a \times n} \tag{B.58} $$

and

- *for all $s \in \mathbb{C}$, a matrix K_b exists to satisfy*

$$ {r_2}^T \mathcal{R}({K_2}^T) \cap \mathcal{N}(P_c) = {r_L}^T \mathcal{R}({K_2}^T) \cap \mathcal{N}(P_c), \tag{B.59} $$

where

$$ K_2 = \begin{bmatrix} K_1 & 0 \\ 0 & K_b \end{bmatrix}, \quad K_b \in \mathbb{R}^{\mathcal{X}_b \times n}, \tag{B.60} $$

and

$$\begin{aligned} r_1 &= \begin{bmatrix} L_0A + MO_\gamma A \\ Y_NA + Y_MO_\gamma A \end{bmatrix}, \\ r_2 &= \begin{bmatrix} L_0(sI - A) + MO_\gamma A \\ Y_N(sI - A) + Y_MO_\gamma A \end{bmatrix}, \\ r_L &= \begin{bmatrix} L_0 \\ Y_N \end{bmatrix}, \\ P_c &= v_2 \begin{bmatrix} O_\gamma A \\ O_\gamma \end{bmatrix}. \end{aligned}$$

Proof. We first consider $R = K_aY_N$ and obtain K_a to satisfy $q_1 = q_2$, from (B.39) and (B.41). Then, we modify $R = \begin{bmatrix} {K_a}^T & {K_b}^T \end{bmatrix}^T$ to also satisfy the other condition – that is, $q_1 = q_3$, from (B.39) and (B.46). The rest of the proof is straightforward, therefore, the details are omitted. □

B.3.1 Design procedure of the functional observer (B.6)

Considering the system (B.1) and the observer (B.6), we can determine the related observer matrices N, J, H, and E as follows:

- Estimate L from (B.56).
- Choose H_1 and E_1 such that $LH_1 = I$ and $LE_1 = 0$ respectively.
- Calculate Λ and Γ from (B.26).
- Choose Z to result in a stable $N = \Lambda - Z\Gamma$ and obtain E and K from (B.24).
- Based on N, E, and K, calculate J from $J = K - NE$.
- Calculate H from (B.15).

B.4 Examples

We consider two examples. In the first example, we use the proposed method to evaluate the existence condition and then design a "0-th derivative available" functional observer for a third-order system. In the second example, we show how the order of a functional observer (designed for the reconstruction of a common functional) decreases when higher derivatives of the input and/or the output signals are available. We then discuss a real-world application of such a functional observer.

B.4.1 Example 1

Consider a third order continuous-time system with the following dynamics,

$$\begin{aligned} A &= \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \quad F = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ C &= \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \quad L_0 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}, \end{aligned} \tag{B.61}$$

and assume that the system has an unknown input (i.e. $\alpha = 1$) and no information about the derivatives of the inputs and outputs is available – that is, $\gamma = 0$ and $\lambda = 0$. Hence, $M_{1,0:\gamma}$, $M_{2,0:\gamma}$ are zero matrices, P is an identity matrix, and $O_\gamma = C$.

For the above system, Theorem 1 in [58] is satisfied. While with

$$rank \begin{bmatrix} CA & CF \\ C & 0 \\ L_0 & 0 \\ Y_1 & 0 \end{bmatrix} = rank \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = 2, \tag{B.62}$$

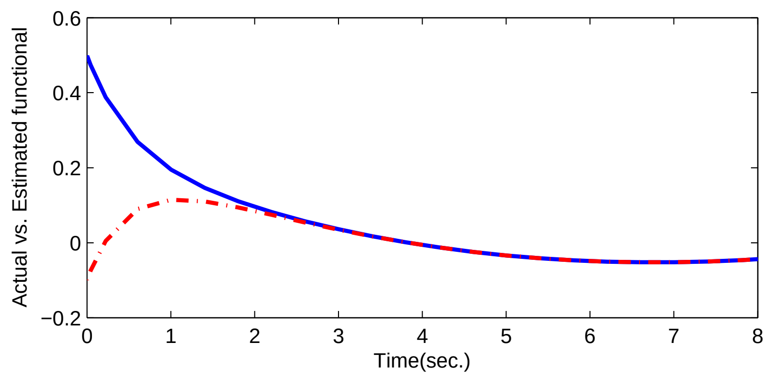
since, $n + rank(CF) = 3$, the first Theorem in [59] (i.e. Theorem 4.6) is not satisfied. Therefore, based on [59], there does not exist a functional observer for this system to reconstruct $L_0 x(t)$.

However, based on Proposition 5 and with having $Y_N x(t)$ to be an empty set, we can show that there exists an unknown-input functional observer with order 1 that can reconstruct $L_0x(t)$.

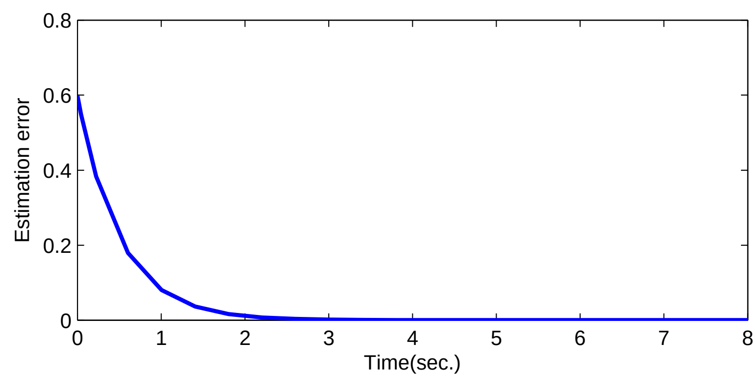
According to Section B.3.1, we can obtain $\Lambda = -0.5$ and $\Gamma = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$. Hence, choosing $Z = [6 \;\; 1]$ will result in a 1st-order observer

$$\begin{aligned} \dot{w} &= -4w - 8y + u, \\ \hat{z} &= w + 3y. \end{aligned} \tag{B.63}$$

Note that the matrix Z is not unique.



(a) The actual functional presented by the solid line vs. the estimated functional presented by the dashed-dotted line



(b) Estimation error of the desired functional

Figure B.1: The error of estimating the functional $z_0(t)$ of (B.62) with the observer (B.63)

Figure B.1 presents the performance of the designed unknown-input functional observer in reconstructing the desired functional, $L_0x(t)$. It is clear that the designed first-order observer (B.63) is capable of estimating $L_0x(t)$ with zero steady-state error.

B.4.2 Example 2

As in [59], we consider the unstable continuous-time unknown-input system, $\alpha = 1$, with dynamics provided in (B.64).

$$
\begin{aligned}
A &= \begin{bmatrix} -0.0226 & -36.617 & -18.897 & -32.09 & 3.2509 & -0.7626 \\ 0.0001 & -1.8997 & 0.9831 & -0.0007 & -0.1708 & -0.005 \\ 0.0123 & 11.72 & -2.6316 & 0.0009 & -31.604 & 22.396 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -30 & 0 \\ 0 & 0 & 0 & 0 & 0 & -30 \end{bmatrix}, \\
B &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 30 \\ 0 & 0 & 0 & 0 & 30 & 0 \end{bmatrix}^T, \quad F = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}^T
\end{aligned}
\tag{B.64}
$$

The output of the system is

$$
C = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{B.65}
$$

and the goal is to reconstruct the functional $L_0x = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} x(t)$. Fernando et.al. [59] showed that there exists a third order observer to reconstruct the desired functional of the states of the mentioned system. Our goal is to investigate whether having a observer with access to the derivatives of the input and output signals can be helpful to reduce the required order of the observer. We, hence, consider an observer with access to the first derivative of the output with $\gamma = 1$ and $\lambda = 0$.

To check the conditions in Proposition 5, we obtain

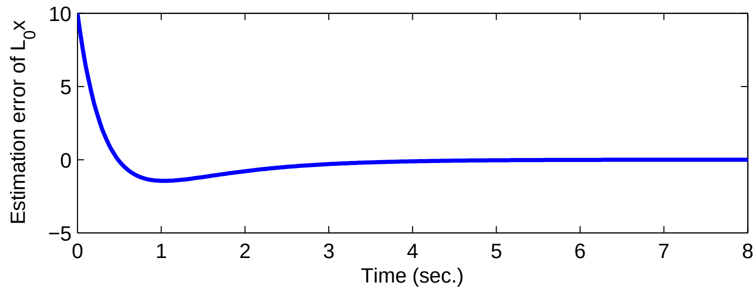
$$\begin{aligned}
Y_N &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0.0006 & -0.0006 \end{bmatrix}, \\
Y_M &= \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}, \\
M &= \begin{bmatrix} 0 & 0 & 11.3766 & 0 \end{bmatrix}, \\
v_2 &= \begin{bmatrix} {e_1}^T & {e_2}^T & {e_4}^T & {e_5}^T & {e_6}^T & {e_7}^T & {e_8}^T \end{bmatrix}^T.
\end{aligned} \tag{B.66}$$

It can be seen that the conditions in Proposition 5 are satisfied. The right hand side and the left hand side of (B.42) are both of rank 6. Besides, (B.43) is not rank deficient for s being selected to be the eigenvalues of $\Lambda|_{L=L_0}$. Hence, there exists a functional observer with access to the first derivative of the output to reconstruct $L_0x(t)$ for this system.

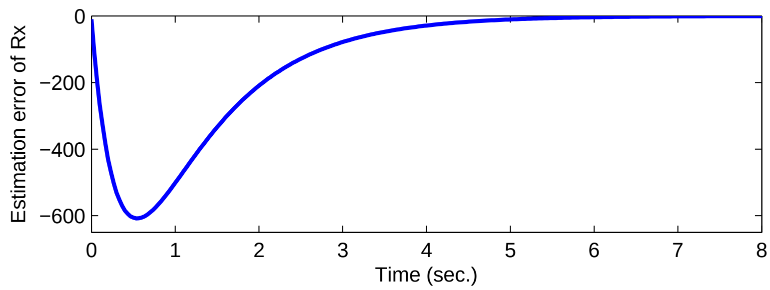
On the other hand, it is straight forward to show that Theorem 9 is not satisfied for $R = \emptyset$, as (B.20) is rank deficient for s being selected to be the eigenvalues of $\Lambda|_{L=L_0}$. Hence, having information about derivatives of outputs, still the desired functional cannot be estimated directly.

To obtain the required order of the functional observer, from (B.50), we obtain $\mathcal{X}_a = 2$ (i.e., at most two additional rows are needed to satisfy (B.19)). However, with Y_N having only one row, the only extension to the desired functional can be $L = \begin{bmatrix} L_0 \\ Y_N \end{bmatrix}$ and we can show that for the extended functional, conditions in Theorem 9 are satisfied. Hence, a second order functional observer with access to the first derivative of the output exists to reconstruct the functional $L_0x(t)$ of the system (B.64).

Using the design procedure provided in B.3.1, we can design a stable functional observer with poles being placed at -3 and -1. We therefore



(a) Estimation error of the desired functional, $L_0x(t)$



(b) Estimation error of the complementary functional, $Rx(t)$

Figure B.2: The error of estimating the functionals of system (B.64) with the observer (B.67)

obtain

$$
\begin{aligned}
N &= \begin{bmatrix} -4.2103 & 0.0125 \\ -311.9616 & 0.2103 \end{bmatrix}, \\
E &= \begin{bmatrix} -0.0061 & -0.0029 & -0.0114 & 0.0111 \\ -2.1213 & 2.2491 & -0.0000 & 0.0130 \end{bmatrix} \times 10^3, \\
J &= \begin{bmatrix} 0.0320 & 0.0445 & 0.0418 & -0.0496 \\ -6.4680 & 1.6618 & 1.4284 & -1.2149 \end{bmatrix} \times 10^3, \\
H &= \begin{bmatrix} -28.3051 & 28.3051 \\ 0.0167 & -0.0167 \end{bmatrix}.
\end{aligned} \tag{B.67}
$$

Having the above observer, the error dynamics in equation (B.9) are modeled. Figures B.2 shows that the estimation error of the desired functional, $L_0x(t)$, and the estimation error of $Rx|_{R=Y_N}$ asymptotically approach zero, hence, the observer makes correct estimations.

Appendix C

List of publications

Journal articles

N. Eskandari, G. A. Dumont, and Z. J. Wang, "Delay-incorporating observability and predictability analysis of safety-critical continuous-time systems," *IET Control Theory and Applications*, vol. 9, pp. 1692–1699(7), July 2015.

N. Eskandari, G. A. Dumont, and Z. J. Wang, "An observer/predictor based model of the user for attaining situation awareness," *IEEE Transactions on Human-Machine Systems*, 2015.

N. Eskandari, Z. J. Wang, and G. A. Dumont, "On the existence and design of functional observers for LTI systems, with application to user modeling," *Asian Journal of Control*, 2014.

N. Eskandari, M. M. Oishi, and Z. J. Wang, "Observability analysis of continuous-time LTI systems with limited derivative data," *Asian Journal of Control*, vol. 16, no. 2, pp. 623–627, 2014.

N. Eskandari, G. A. Dumont, and Z. J. Wang. 'Bounded-error delayed functional observer/predictor: Existence and design," *submitted*.

Conference papers

N. Eskandari, Z. J. Wang, and G. Dumont, "Modeling the user as an observer to determine display information requirements," in *In Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, Manchester, UK, October 2013, pp. 267–272.

N. Eskandari, and M. Oishi, "Computing observable and predictable subspaces to evaluate user-interfaces of LTI systems under shared control," in *Proceedings of the IEEE conference on Systems, Man and Cybernetics*, Alaska, USA, October 2011, pp. 2803–2808.

Book Chapters

R. Cortez, D. Tolic, I. Palunko, **N. Eskandari**, R. Fierro, M. Oishi, and J. Wood, A hybrid framework for user-guided prioritized search and adaptive tracking of maneuvering targets via cooperative UAVs, In *Advances in Intelligent and Autonomous Aerospace Systems*, in the *Progress in Aeronautics and Astronautics Series* American Institute of Aeronautics and Astronautics, Reston, VA, vol. 241, 2012