Estimation d'état et de paramètres de modèles de masses neuronales : une approche par la théorie du contrôle

Adel Malik Annabi

Centre Inria d'Université Côte d'Azur, équipe McTAO

Présentée en vue de l'obtention du grade de docteur en Mathématiques
d'Université Côte d'Azur
Dirigée par : Jean-Baptiste Pomet.
Á soutenir le : 02/07/2026

Devant le jury, composé de :
Pauline Bernard, Maitre-Assistant, Mines Paris
Eric Busvelle, Professeur, Université Toulon
Mathieu Desroches, Directeur de recherche, Inria Montpellier
Alessio Franci, Professeur assistant, Université de Liège
Romain Postoyan, Directeur de recherche, CNRS Université de Lorraine
Dario Prandi, Directeur de recherche, CNRS Université Paris Saclay
Sophie Tarbouriech, Directeur de recherche, CNRS INSA Toulouse
Ludovic Sacchelli, Chargé de recherche, Inria d'Université Côte d'Azur

Estimation d'état et de paramètres de modèles de masses neuronales : une approche par la théorie du contrôle

State and parameter estimation of neural mass models: a control-theoretic approach

Jury :

Directeur de thèse :
Jean-Baptiste Pomet, Directeur de recherche, Inria d'université Côte d'Azur

Président du jury :

Rapporteurs :
Romain Postoyan, Directeur de recherche, CNRS Université de Lorraine
Alessio Franci, Professeur assistant, Université de Liège

Examinateurs :
Pauline Bernard, Maitre-Assistant, Mines Paris
Mathieu Desroches, Directeur de recherche, Inria Montpellier
Eric Busvelle, Professeur, Université Toulon
Sophie Tarbouriech, Directeur de recherche, CNRS INSA Toulouse

Co-encadrants :
Dario Prandi, Directeur de recherche, CNRS Université Paris Saclay
Ludovic Sacchelli, Chargé de recherche, Inria d'université Côte d'Azur

Estimation d'état et de paramètres de modèles de masses neuronales : une approche par la théorie du contrôle

Résumé

Cette thèse porte sur le problème de l'estimation de l'état et des paramètres de modèles de masse neuronale à partir de mesures partielles. Nous nous concentrons sur des réseaux de systèmes dynamiques couplés avec des non-linéarités saturantes. Ces modèles, dérivés de moyennes collectives sur de grandes populations neuronales, sont largement utilisés pour décrire l'activité neuronale à l'échelle mésoscopique et macroscopique et constituent la base d'applications allant de la stimulation cérébrale aux interfaces cerveau-ordinateur. Cependant, ils souffrent de deux difficultés structurelles : une forte incertitude sur les paramètres et la difficulté de mesurer l'état complet. La thèse est divisée en deux parties.

La première traite de l'estimation d'état pour les modèles de masse neuronale et de champ neuronal. Après avoir passé en revue certaines conceptions d'observateurs connues pour les modèles de masse neuronale, une nouvelle conception d'observateur est proposée, qui exploite une reconstruction en temps fini d'une sortie non linéaire via un différentiateur exact. Le terme de correction résultant découple une partie du couplage non linéaire dans la dynamique de l'erreur et fournit une condition LMI moins conservative que les conditions existantes ; des simulations numériques sont présentées sur le problème de reconstruction de champ à partir d'un nombre fini de canaux de mesure. Deux articles inclus traitent de l'estimation d'état pour un modèle de champ neuronal de faible dimension du cortex visuel primaire, pour lequel seule une moyenne spatiale du champ est disponible. Le premier construit un observateur hybride à grand gain, caractérise les singularités d'observabilité du modèle et prouve la convergence pratique sous une condition d'excitation persistante. Le second propose une nouvelle transformation qui plonge le système dans une forme triangulaire spécifique sur laquelle un observateur homogène à modes glissants peut être conçu, permettant une reconstruction en temps fini de l'état. Les singularités d'observabilité sont à nouveau traitées via un mécanisme de commutation, et des simulations numériques sont fournies.

La deuxième partie traite de l'estimation de paramètres. Dans nos modèles d'intérêt, les paramètres inconnus apparaissent dans le modèle multipliés par une non-linéarité saturante appliquée à un état non mesuré. Le couplage est bilinéaire et rend le problème inverse mal posé. Nous proposons de nouveaux schémas d'estimation qui incluent un control dont le but est de conduire l'état inconnu vers différentes régions où la saturation atteindra un plateau à une valeur connue, permettant ainsi la récupération des paramètres. Cela est possible malgré l'incertitude sur l'état et les paramètres, car la structure de la saturation est connue. Une fois les paramètres récupérés, nous présentons l'estimation d'état. Nos schémas fournissent une estimation conjointe de l'état et des paramètres en temps fini. Nous développons d'abord cette méthode sur un système à deux dimensions, où le cas linéaire est montré non identifiable tandis que le cas saturant ne l'est pas. Nous étendons le schéma au cas de dimension finie, en définissant le type de systèmes adéquat sur lequel cette méthode peut être appliquée, puis au cas de dimension infinie où l'estimation exacte est relaxée à toute précision souhaitée. Ces résultats sont combinés en un protocole d'estimation complet pour un réseau à deux couches

de systèmes de dimension finie, où nous montrons que nous pouvons récupérer tous les paramètres inconnus apparaissant dans la dynamique des populations mesurées, l'état de la population non mesurée et la plupart de ses paramètres.

Mots clés : estimation d'état, synthèse d'observateurs, estimation de paramètres, identification non linéaire, systèmes non linéaires, théorie du contrôle, modèles de masse neuronale, modèles de champ moyen, champs neuronaux, neurosciences.

State and parameter estimation of neural mass models: a control-theoretic approach

Abstract

This thesis addresses the problem of estimating the state and parameters of neural mass models from partial measurements. We focus on networks of coupled dynamical systems with saturating nonlinearities. These models, derived from collective averages over large neuronal populations, are widely used to describe neural activity at a mesoscopic and macroscopic scale and serve as the basis for applications ranging from brain stimulation to brain-computer interfaces. However, they suffer from two structural difficulties: high parameter uncertainty and the difficulty of measuring the full state. The thesis is divided into two parts.

The first one addresses state estimation for neural mass and neural field models. After reviewing some known observer designs for neural mass models, a new observer design is proposed that exploits a finite-time reconstruction of a nonlinear output via an exact differentiator. The resulting correction term decouples part of the nonlinear coupling in the error dynamics and yields an LMI condition that is less conservative than existing ones; numerical simulations are provided on the problem of field reconstruction from a finite number of measurement channels. Two included articles treat state estimation for a low-dimensional neural field model of the primary visual cortex, for which only a spatial average of the field is available. The first one constructs a hybrid high-gain observer, characterizes the observability singularities of the model, and proves practical convergence under a persistent excitation condition. The second one proposes a new transformation that embeds the system into a specific triangular form on which a homogeneous sliding mode observer can be designed, achieving finite-time reconstruction of the state. Observability singularities are again handled via a switching mechanism, and numerical simulations are provided.

The second part addresses parameter estimation. In our models of interest, the unknown parameters enter the model multiplied by a saturating nonlinearity applied to an unmeasured state. The coupling is bilinear and renders the inverse problem ill-posed. We propose novel estimation schemes that include a control, whose sole purpose is to drive the unknown state to various regions where the saturation will plateau at a known value which enables the recovery of parameters. This is possible despite state and parameter uncertainty because the saturation structure is known. Once the parameters are retrieved, we showcase state estimation. Our schemes provide joint estimation of state and parameters in finite time. We first develop this method on a two-dimensional system, where crucially the linear case is shown to be unidentifiable while the saturating case is not. We extend the scheme to the finite dimensional case, where we define the adequate type of systems on which this method can be applied, and then further extend to the infinite-dimensional case where the exact estimation is relaxed up to any desired accuracy. These results are combined into a complete estimation protocol for a two-layer network of finite-dimensional systems, where we show that we can retrieve all the unknown parameters appearing in the dynamics of measured populations, the state of the unmeasured population and most of its parameters.

Keywords: state estimation, observer design, parameter estimation, nonlinear identification, nonlinear systems, control theory, neural mass models, mean-field models, neural fields, neuroscience.

TABLE OF CONTENTS

Part I

GENERAL INTRODUCTION

1

INTRODUCTION

1.1 INTRODUCTION TO ESTIMATION IN DYNAMICAL SYSTEMS

Dynamical systems provide a powerful mathematical framework for modeling the evolution of phenomena across time and space, capturing the essential mechanisms that govern natural and engineered processes. From the orbits of celestial bodies to biological systems, these systems encode the fundamental principle that the present state determines the future. In the realm of neuroscience, this approach becomes particularly compelling, as the brain itself can be understood as a complex, evolving system whose dynamics underpin cognition, perception, and behavior [36, 60]. The challenge, and the art, lie in formulating mathematical descriptions that are both sufficiently rich to capture the brain's complexity and sufficiently tractable to permit analysis.

The brain is composed of neurons, specialized cells that integrate electrochemical inputs and transmit signals through synaptic connections [109]. The seminal work by Hodgkin-Huxley [100] provided a detailed biophysical model of neuronal action potential generation. This model remains the de facto canonical description of the single neuron. However, the brain is a multi-scale system [36, 60], where phenomena at the level of individual neurons [100] (the microscopic scale) give rise to emergent patterns of activity at the level of sufficiently large populations and cortical regions (mesoscopic scale [167, 168]) which then further interact to produce the large-scale dynamics observed at the level of the whole brain [107, 108] (macroscopic scale). Each of these scales offers a different window into brain function, and each comes with its own modeling challenges.

The human brain contains on average 86 billion neurons [15], each forming thousands of synaptic contacts. Modelling the brain at the level of individual neurons is computationally intractable. Moreover, many phenomena such as movement, perception, synchrony, wave propagation are emerging from the coordinated interactions of large populations. Macroscopic observables such as the EEG or the BOLD signal measured by fMRI are inherently aggregated quantities that reflect population-level dynamics rather than single-neuron activity [36]. Neural Mass Models (NMMs) and their spatially extended counterparts are mathematical frameworks designed to bridge the gap [36, 168, 106]. Rather than capturing the dynamics of individual neurons, they focus on the collective behavior of large neuronal populations while abstracting away the microscopic details. By averaging the activity of many neurons, these models can capture the emergent dynamics that arise from their interactions at the mesoscopic scale [36, 168, 167, 106]. Neural mass models can then be interconnected to form large-scale brain network models that capture the interactions between different brain regions, providing a powerful tool for understanding the macroscopic dynamics of the brain [107, 108, 60, 142, 46]. These discrete population models can be further extended to a continuum limit, giving rise to neural field models [3, 107, 52], where the state variable becomes a spatially distributed field evolving over a cortical manifold.

Many of these models of brain activity can be cast into the following general form [168, 167, 106, 56, 166, 165, 92, 3, 52]

$$\begin{cases} \dot{V} = LV + WS(V) + Bu, \\ y = h(V). \end{cases} \tag{1.1}$$

While the precise interpretation of each term varies across models, in the case of neural fields [3, 107, 52] the terms take the following meaning: The state vector V, belonging to a function space over the cortical domain, represents the dynamic variables of the neural populations, such as their mean firing rates. The operator L represents the intrinsic decay or damping dynamics of each isolated population. The operator W is the core of the network model, encoding the strength and topology of the synaptic interconnections between different neuronal units. The nonlinear function $S(\cdot)$, often chosen to be a sigmoid, acts component-wise on the states V and represents the transformation of the mean membrane potential of a population into its mean firing rate, thereby introducing a fundamental and biologically plausible saturation effect. The control input u represents an external stimulus, which can be non-invasive brain stimulation such as sensory stimulation, or invasive as deep stimulation on certain parts or regions of the brain. The operator B shapes how the external input enters the system.

A central challenge in both clinical and experimental neuroscience is the *observation problem*, which is explicitly represented in the model by the output equation $y = h(V)$. Most of the time we will consider $y = CV$ in voltage, or $y = CS(V)$ in activity (spike per second). In practice, we never have access to the full state vector V; our measurements are inherently partial and indirect. In the case of EEG, for instance, the matrix C models how the cortical source activities (the state V) are projected onto a limited set of scalp electrodes to produce the measured signals y. In the case of Local Field Potential (LFP) recordings, the measurement operator C acts as a convolution of the field with a spatial kernel at the measurement point, the kernel is usually taken to be Gaussian [83, 105] (see Section 3.2 for more details). In either case, this operator is severely ill-conditioned, meaning vastly different underlying brain states can produce nearly identical sensor readings. This problem of inferring the rich internal state from a limited, often blurred, external observation is not merely a technical nuisance; it is a fundamental feature of brain dynamics that drives the development of sophisticated state estimation and signal processing techniques.

Closely related and equally compelling is the *identification problem*. Given a measured output y, can we identify the unknown parameters of the model that generated it? The prime target of such an inquiry is often the operator W, as it embodies the functional architecture of the brain network. Discerning the structure of W is of paramount importance for understanding healthy brain function and diagnosing neurological disorders, where specific connectopathies are believed to be a root cause of such disorders. However, this inverse problem is fraught with ambiguity. The brain is a complex, noisy, and high-dimensional system where different combinations of parameters can yield nearly identical observed dynamics. This parameter ambiguity, a manifestation of the broader issue of identifiability in nonlinear systems, means that uniquely determining the brain's "wiring diagram" from data alone is an ill-posed problem. It necessitates the integration of prior knowledge, the development of specialized estimation algorithms, and a careful analysis of which network features can be reliably inferred from the available partial measurements. Solving this puzzle is a critical step towards creating personalized, model-based diagnostics and therapies for brain disorders.

Given that the brain is a dynamical nonlinear system whose parameters may themselves evolve at different time-scales [36] than the neural states, there is an imperative to develop methods for *online* estimation. This requires estimating the system's state and parameters in real-time as data is being acquired, rather than through offline batch processing. Such capabilities are crucial for developing adaptive brain-computer interfaces, closed-loop neuromodulation therapies, and for understanding how neural circuits reconfigure during learning or in response to neurological disorders.

To solve these problems, one may conceptualize the brain as an input-output system, where the input u represents an external stimulus and the output y corresponds to partial measurements acquired via modalities such as EEG, ECoG, or MEG. This abstraction allows us to formalize the problems of state and parameter recovery from a control-theoretic viewpoint.

Let us present this framework from a more theoretical point of view. Consider the following nonlinear system

$$\dot{x}(t) = f(x(t), u(t), \theta) \qquad \text{(State Equation)} \qquad (1.2)$$

$$y(t) = h(x(t), u(t), \theta) \qquad \text{(Output Equation)} \qquad (1.3)$$

where x denotes the state vector whose values are in $\mathbb{R}^n$. We denote the set of admissible initial conditions by $\mathcal{X}_0 \subset \mathbb{R}^n$. The variable u represents the input, assumed to be at least measurable and locally bounded, taking values in $\mathbb{R}^m$. The variable $\theta \in \mathbb{R}^p$ constitutes the vector of unknown dynamical parameters, and $y \in \mathbb{R}^q$ corresponds to the measured output.

The functions f and h are assumed known, while parameters θ must be estimated from experimental data.

The fundamental challenges of state and parameter recovery can be formally cast as two distinct mathematical problems.

Observation/State Estimation Problem: Construct an estimate $\hat{x}$ of the state trajectory x of (1.2) such that for any admissible input u and any initial condition $x_0 \in \mathcal{X}_0$, using only the knowledge of input and output histories $u_{[0,t]}$ and $y_{[0,t]}$, the estimate $\hat{x}$ asymptotically approaches x.

The observation problem can be equivalently formulated as: given the input signal, determine the unique initial condition that could have generated the observed output trajectory. This reformulation immediately raises the question of uniqueness - whether distinct initial states can produce identical input-output pairs - which relates directly to the fundamental concept of observability that will be rigorously addressed subsequently.

The observation problem is fundamentally an *inverse problem*: given the known inputs u and measured outputs y of a dynamical system, reconstruct the hidden internal state x that generated them. This inverse problem is inherently challenging due to its ill-posed nature, i.e., multiple state trajectories can potentially explain the same input-output data, particularly when measurements are partial and noisy.

The notion of observability provides a necessary and sufficient condition to ensure the well-posedness of this inverse problem. Formally, the system (1.2) is said to be observable if for every pair of distinct initial states $(x_0, \tilde{x}_0) \in \mathcal{X}_0 \times \mathcal{X}_0$ with output $(y, \tilde{y})$, there exists a finite time $T \geq 0$ and an admissible input $u \in \mathcal{U}$ such that the corresponding output trajectories are distinguishable:

$$x_0 \neq \tilde{x}_0 \quad \Rightarrow \quad \exists\, t \in [0, T], \quad y(t) \neq \tilde{y}(t).$$

Equivalently, observability requires that the mapping from initial conditions to output trajectories is injective:

$$\forall t \geq 0 \quad y(t) = \tilde{y}(t) \quad \Rightarrow \quad x_0 = \tilde{x}_0.$$

This condition ensures that no two different initial states can produce identical input-output pairs, thereby guaranteeing that the state reconstruction problem admits a unique solution.

Historically, two natural approaches have been explored to solve this inverse problem in an online context. The first consists of simulating the system dynamics in parallel for a set of candidate initial conditions, progressively eliminating those whose output trajectories deviate significantly from the measurements. While conceptually straightforward, this method suffers from exponential computational complexity and extreme sensitivity to model inaccuracies. The second approach formulates the problem as the minimization of a cost function measuring the discrepancy between predicted and measured outputs:

$$\hat{x}(t) \in \mathrm{Argmin}_{\hat{x}} \int_0^t \|h(\hat{x}, u(\tau), \theta) - y(\tau)\|^2 d\tau. \tag{1.4}$$

However, this optimization-based framework faces significant challenges for real-time implementation, including non-convexity, the presence of local minima, and substantial computational demands that make it unsuitable for online applications.

We adopt an alternative perspective rooted in dynamical systems theory: we reformulate the inverse problem as the design of an *observer*, i.e., an auxiliary dynamical system that solves the state reconstruction problem recursively and in real-time. Rather than treating state estimation as a batch optimization or parallel simulation problem, the observer approach provides a continuous, feedback-based solution that dynamically corrects its estimates using the measured output y and input u. This approach consists in constructing an auxiliary dynamical system, driven by the same input u and the measured output y of the original system, whose output or a transformation of it, provides an asymptotic estimate $\hat{x}(t)$ of the true state $x(t)$. Formally, such a non-linear observer for (1.2) takes the structure:

$$\begin{cases} \dot{\hat{z}} = \mathcal{F}(\hat{z}, y, u), \\ \hat{x} = \mathcal{T}(z, y, u), \end{cases} \tag{1.5}$$

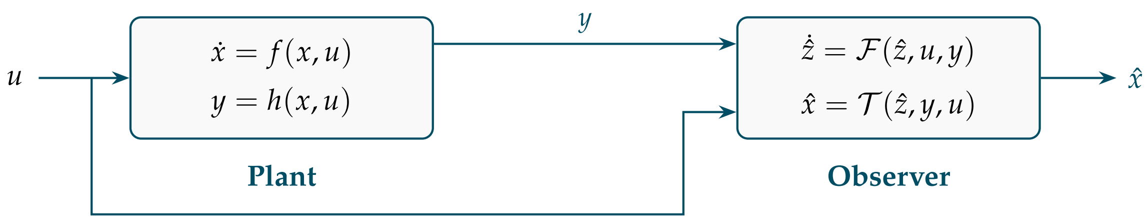


Figure 1.1: A schematic of the observer design problem. The plant is described by a nonlinear system with state x, input u, and output y. The observer generates an estimate $\hat{x}$ of the state using the input and output measurements. The design goal is to ensure that $\hat{x}$ converges to x over time.

where $\hat{z}$ and $\hat{x}$ take elements in $\mathbb{R}^{d_z}$ and $\mathbb{R}^{d_x}$ respectively, $\mathcal{F}$ and $\mathcal{T}$ are a family of continuous maps indexed by the control u. We denote $\mathcal{Z}_0$ the set of admissible initial conditions $\hat{z}_0$.

Ideally, we wish the following property: For any initial condition $\hat{z}_0 \in \mathcal{Z}_0 \subset \mathbb{R}^{d_z}$, $x_0 \in \mathcal{X}_0 \subset \mathbb{R}^{d_x}$, for any admissible control u, solutions of the couple ((1.2)-(1.5)) verify

$$\lim_{t \to +\infty} |\hat{x}(t) - x(t)| = 0. \tag{1.6}$$

This paradigm offers several distinct advantages over the methods previously discussed. Firstly, it provides a recursive and computationally efficient framework for online state estimation, avoiding the exponential complexity of simulating ensembles of trajectories or solving non-convex optimization problems at each time step. Secondly, it establishes a rigorous theoretical foundation based on Lyapunov stability theory, enabling formal guarantees on estimation error convergence and robustness properties. Thirdly, the observer structure naturally integrates with control design, facilitating the development of output-feedback strategies, and elegantly extends to the simultaneous estimation of states and unknown parameters through adaptive observer frameworks. This makes it particularly suitable for real-time applications in neuroscience where computational resources are constrained and theoretical guarantees are desirable.

Returning to the parameter estimation, we can define the identification problem as follows

Identification/Parameter Estimation Problem: Design an admissible input $u \in \mathcal{U}$ such that for any initial condition $x_0 \in \mathcal{X}_0$, we can determine the unknown parameter vector θ from the input-output data $u_{[0,t]}$ and $y_{[0,t]}$. Mainly, construct an estimate $\hat{\theta}(t)$ that asymptotically approaches the true parameter θ on $[0, +\infty)$. The system is identifiable if for any two parameter values $\theta_1, \theta_2 \in \Theta$, if there exist $u \in \mathcal{U}$ such that the equality of output maps $y(t, \theta_1) \equiv y(t, \theta_2)$ for all $t \geq t_0$ implies $\theta_1 = \theta_2$.

The structural identifiability of the parameter vector θ hinges on the injectivity of the parameter-to-output map. The identification problem can be equivalently formulated as: determine whether the model parameterization θ is unique given the input-output behavior. This reformulation immediately raises the question of distinguishability: whether different parameter values can generate identical input-output pairs for all admissible inputs. This question is particularly important for neural systems, where parameter redundancies and structural symmetries often lead to multiple parameter configurations explaining the same observed dynamics. In the neuro-computational community, this phenomenon is called *degeneracy/neurodegeneracy* [2] in certain contexts.

The parameter identification problem has been extensively studied within the system identification literature, covering both offline and recursive approaches [121, 147, 93, 130]. However, in this work, we will specifically address this problem through the lens of observer theory, which provides a dynamical systems framework for real-time parameter estimation. From this perspective, two principal methodologies emerge for joint state and parameter estimation [29].

The first approach is state extension, where unknown parameters are treated as additional state variables with zero dynamics, leading to an extended system of (1.2):

$$\begin{aligned} \dot{x}_e(t) &= \begin{bmatrix} f(x(t), u(t), \theta(t)) \\ 0 \end{bmatrix} \\ y(t) &= h(x(t), u(t), \theta(t)) \end{aligned}$$

where $x_e(t) = [x(t), \theta(t)]^T$. A nonlinear observer can then be designed for this extended system, though this approach may destroy beneficial structural properties of the original system.

The second approach employs adaptive observers, which maintain the separation between state and parameter dynamics. The adaptive observer approach relies on a known observer design for the system when all parameters are known. This observer provides an estimate $\hat{x}_\theta$ for x when θ is known, together with a Lyapunov function V for the observation error dynamics $\hat{x}_\theta - x$. When parameters are unknown, the observer uses parameter estimates $\hat{\theta}$ and incorporates an adaptation law on an estimate $\hat{\theta}$ designed to preserve the stability properties of the original error dynamics through Lyapunov-based synthesis methods. Under some adequate excitation condition, convergence of the estimate $\hat{\theta} \to \theta$ may be proven alongside the convergence of $\hat{x}$.

Adaptive observers have been overwhelmingly studied in the case where the parameter θ in (1.2) appears linearly. They are rooted in the Recursive Least Squares (RLS) algorithm [121, 172] which allows one to estimate a constant or very slowly evolving parameter from state measurements. Most results on adaptive observers are obtained when the structure of system (1.2) allows for the estimation of x independently of θ [29, 171]. The estimate of the state is then used to construct an estimate of the parameters as it was done in [40, 41, 44]. Another popular form is when the system is in a cascade form which allows the sequential estimation of the state and parameters [29, 77, 80]. Although this is not always the case as in [137].

In conclusion, control theory may be used in adequate cases to jointly estimate the state x and parameter θ of a dynamical system (1.2) in real time, through the specific design of an input u and the knowledge of the measurement y.

1.2 STRUCTURE OF THE THESIS

The thesis is organized in three main parts. The first one covers background and context. The second is devoted to state estimation, and the third to parameter estimation.

Part one is made of the present chapter and the next one that introduces the neuroscience setting in a bit more detail. It traces the path from individual neurons to large-scale brain network models and reviews the experimental and modeling challenges that motivate the estimation problems studied here. It also discusses why the control-theoretic perspective differs from the data assimilation approaches that are standard in computational neuroscience.

The following two chapters are the beginning of part 2, they study the state estimation problem for some neural mass models, especially the Wilson-Cowan model and those obtained from a finite-dimensional reduction of neural fields. The first one reviews existing observer designs for this class which includes: copy-of-the-system observers under global contraction, KKL and cascade sliding-mode designs exploiting triangular connectivity, interval observers under monotonicity, and Luenberger designs based on circle-criterion LMI conditions. For each approach, the structural assumptions required for convergence are made explicit. Chapter 4 then proposes a new observer design. The key idea is to reconstruct a nonlinear output in finite time via a super-twisting differentiator, then use it to inject a correction that reduces the effective coupling in the error dynamics. The resulting LMI condition is less conservative than the standard one, with gains that remain bounded as the coupling strength grows.

Chapters 5 and 6 are two published articles on state estimation for a low-dimensional neural field model of the primary visual cortex, where only a spatial average of the field is available. The first characterizes the observability singularities of the system and constructs a hybrid high-gain observer by embedding the system into a triangular form when it is beyond the singular regions. The second article improves on this approach by proposing a new transformation that embeds the system into a triangular form where a finite-dimensional homogeneous observer is possible and achieves finite-time convergence. The stability of this observer is also discussed as well as the switching mechanism.

These two articles rely on a specific kind of observer. A brief overview of these observers is provided in the appendix. The reader familiar with these methods can proceed directly; the appendix is meant as a self-contained reference rather than introductory reading.

The third part of the thesis addresses parameter identification. We wish to estimate parameters in highly uncertain systems that exhibit a certain structure: An unknown parameter to retrieve is multiplied by a saturating nonlinearity applied to an unmeasured state. This structure is common in neural mass models. Throughout the whole chapter, we develop the necessary tools to solve this problem, which is ill-posed in the linear case, and then apply it to the identification of saturated dynamical networks.

Chapter 7 works through the simplest case, a two-dimensional system with one measured state, and shows precisely why classical methods fail. It proposes a novel approach that exploits the very nature of nonlinearity. By steering the hidden state into the upper and lower saturation plateaus via open-loop control, the unknown gain and unknown state decouple and can be read off from the measured output. This converts a seemingly ill-posed problem into a finite-time estimation procedure.

Chapter 8 extends this to networks of arbitrary dimension with measurements from a single layer, developing a general canonical form for parameter estimation. It handles simultaneous estimation of states and parameters through a structured sequence of steering steps and averaging estimators. Chapter 9 carries the same framework to the infinite-dimensional setting.

Chapter 10 presents the full identification protocol for saturated dynamical network systems. The interconnection matrices are recovered layer by layer, under the assumption that only the first layer is measured, combining the estimators from the preceding chapters into a single procedure with finite-time guarantees.

Chapter 11 is the article, published at IEEE CDC 2025 conference, containing the results on the identification of hierarchical networks together with numerical illustrations. This article was the sketch of the full parameter estimation procedure, which was then extended and refined in the previous chapters. Although this chapter comes last, we stress that the results it contains were obtained before the more general results of the previous chapters, and that the latter were developed in an attempt to extend and refine the method proposed in this article. It is therefore the most "crude" version of the results.

2

NEUROSCIENCE CHALLENGES IN IDENTIFICATION AND OBSERVATION

2.1 FROM NEURONS TO LARGE SCALE MODELS OF THE BRAIN

2.1.1 *From single neurons to population dynamics*

The human brain operates as a complex multiscale dynamical system where movement, cognition, and perception emerge from the collective activity of neurons organized within cortical circuits and large-scale systems [58]. While single neuron spiking was adequately captured by the Hodgkin-Huxley model [101], understanding how collective neural behavior arises at macroscopic scales remains a central challenge in neuroscience.

At the microscopic level, neuronal activity is governed by nonlinear differential equations describing ion flow across cell membranes. The Hodgkin-Huxley model describes the membrane potential through four nonlinear ODEs that capture how voltage-gated sodium and potassium channels generate action potentials [100]. While bio-physically precise, simulating large networks of such detailed neurons becomes computationally prohibitive, motivating the development of population-level approaches.

The transition to mesoscopic scales requires a shift from tracking individual spikes to characterizing population-level activity. This challenge has led to the development of neural mass models, which describe the collective behavior of neuronal populations rather than individual neurons.

2.1.2 *Neural mass models: Population-level dynamics*

Neural mass models (NMMs) provide a phenomenological approach to modeling the average activity of neuronal populations [82, 36]. These models operate at the mesoscopic scale, typically representing cortical columns containing thousands to hundreds of thousands of neurons. Rather than simulating individual spikes, NMMs describe how population-level variables such as average firing rates or synaptic activities evolve over time.

The fundamental assumption underlying neural mass models is that the collective behavior of neuronal populations can be captured by a small number of state variables and differential equations. This dimensional reduction makes large-scale brain simulations computationally feasible while preserving essential features of neural dynamics such as oscillations, synchronization, and state transitions. A classic example of a neural mass model is the Wilson-Cowan equations [168], which describe the interaction between excitatory (E) and inhibitory (I) neuronal populations. The original equations are:

$$\tau_e \frac{dE}{dt} = -E + S_e(w_{ee}E - w_{ei}I + P(t)), \tag{2.1}$$
$$\tau_i \frac{dI}{dt} = -I + S_i(w_{ie}E - w_{ii}I + Q(t)). \tag{2.2}$$

Here, $E(t)$ and $I(t)$ represent the firing rates of excitatory and inhibitory populations, τ_e and τ_i are time constants, w_{jk} are synaptic coupling strengths, and $P(t)$, $Q(t)$ are external inputs. The sigmoidal functions $S_e(x)$ and $S_i(x)$ map the net input to each population's output firing rate. Through a change of variables where $V_E := w_{ee}E - w_{ei}I + P(t)$ and $V_I := w_{ie}E - w_{ii}I + Q(t)$ represent the post-synaptic membrane potentials of the excitatory and inhibitory populations respectively, the Wilson-Cowan equations can be rewritten as:

$$\tau_e \frac{dV_E}{dt} = -V_E + w_{ee}S_E(V_E) - w_{ei}S_I(V_I) + \tilde{P}(t), \tag{2.3}$$
$$\tau_i \frac{dV_I}{dt} = -V_I + w_{ie}S_E(V_E) - w_{ii}S_I(V_I) + \tilde{Q}(t). \tag{2.4}$$

Here $\tilde{P}(t)$ and $\tilde{Q}(t)$ represent the transformed external inputs. In this formulation, the variables V_E and V_I represent the post-synaptic membrane potentials through an appropriate change of scale, and the parameters appear outside the sigmoid functions, making the network structure more explicit. This representation highlights the recurrent connectivity between populations and facilitates the extension to large-scale networks. Should the model be extended to include more populations, the equations can be written in a more compact block form as:

$$\dot{V} = -DV + \begin{pmatrix} W_{ee} & W_{ei} \\ W_{ie} & W_{ii} \end{pmatrix} S(V) + Bu, \tag{2.5}$$

where $S(V) = (S_i(V_i))_{i=1}^n$ is the vector of nonlinearities, D is a diagonal matrix of decay rates, W is the synaptic connectivity matrix, and Bu represents external inputs.

In fact, most neural mass models can be cast into a general form which can be seen as a Lur'e system [115]:

$$\dot{V} = AV + Wf(V) + u(t), \tag{2.6}$$

where $V(t) \in \mathbb{R}^n$ is the vector of mean membrane potentials of the n populations, $A \in \mathbb{R}^{n\times n}$ and $W \in \mathbb{R}^{n\times n}$, $f : \mathbb{R}^n \to \mathbb{R}^n$ is a particular nonlinear function representing the population response, and $u(t) \in \mathbb{R}^n$ collects external inputs. The signification and structure of A,W and f depend on the specific model, but the general form is widely used in the literature. Indeed, models such as the Wilson-Cowan [168, 167], Jansen-Rit [106] are some of the most well-known examples, further models proposed in the literature including [166, 165, 56, 92] can also be cast into this general form. This formulation makes the linear-plus-nonlinear structure of the dynamics explicit, which is particularly convenient for observer and identification design [48, 31].

It makes sense therefore, to study the estimation problem for this general class of models, as it will cover a wide range of models used in the neuroscience literature. In particular, we mainly focus in this work on equations (2.6) that exhibit saturation effects. This is the case in the third part of the thesis, where the very nature of the saturation is exploited to achieve parameter estimation.

We mainly focus on this specific case because it is the usual structure for a general Wilson-Cowan model, and is also the form obtained when reducing a *neural fields model* (that will be introduced in 2.1.4) to a finite-dimensional system, as will be shown in 3.2.

2.1.3 *Mean-field models: Statistical foundations*

While neural mass models are often formulated phenomenologically, mean-field theory provides a rigorous mathematical foundation for deriving population-level equations from networks of spiking neurons [42]. The mean-field approach begins with a network of individual spiking neurons with specific connectivity patterns and derives equations for the collective dynamics through statistical averaging.

The derivation typically involves several key steps: first, specifying the microscopic dynamics of individual neurons and their synaptic interactions; second, defining the network architecture and coupling between neurons; third, applying statistical mechanics principles to derive equations for population averages; and finally, making simplifying assumptions such as the diffusion approximation to obtain tractable equations.

The Fokker-Planck formalism provides a powerful framework for this derivation, describing how the probability distribution of neuronal states evolves over time. Under appropriate assumptions about network homogeneity and weak correlations, the complex dynamics of thousands of interacting neurons can be reduced to a small set of differential equations for population means and variances.

The Wilson-Cowan model can be derived through such mean-field approaches by starting from networks of integrate-and-fire neurons with specific connectivity patterns and applying statistical averaging techniques [1]. This derivation provides a clearer connection between the phenomenological parameters of the Wilson-Cowan equations and the biophysical properties of individual neurons and synapses.

2.1.4 *Neural field models: Spatial continuum extension*

While neural mass models describe localized populations, many neural phenomena involve wave propagation and spatial pattern formation across continuous cortical sheets. Neural field models extend the population approach to spatial domains by treating the cortex as a continuous excitable medium [4, 53]. The transition from discrete neural mass models to continuous neural field models can be achieved by considering the continuum limit of coupled Wilson-Cowan equations. This leads to the Amari [3, 167] neural field equation describing the evolution in a slice of cortex $\Omega \subset \mathbb{R}^2$:

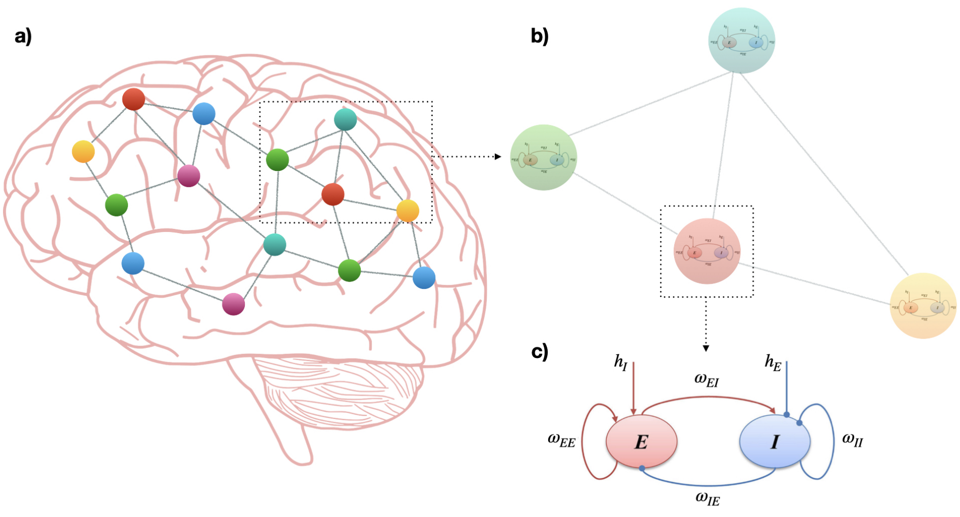


Figure 2.1: From [123]. This figure illustrates a Neural Mass Network Model. a) Nodes represent a population of neurons. b) Within each region, or node of the network, there are interacting subpopulations of excitatory and inhibitory cells, in accord with the standard Wilson-Cowan model. c) Standard Wilson-Cowan model.

$$\tau \frac{\partial u(x,t)}{\partial t} = -u(x,t) + \int_{\Omega} w(|x-y|) S(u(y,t)) dy + I(x,t). \tag{2.7}$$

Here, $u(x,t)$ represents the neural activity at cortical position x and time t, $w(|x-y|)$ is a synaptic kernel defining the strength of connections between points x and y, S is a sigmoidal activation function, and $I(x,t)$ represents external inputs. The synaptic kernel $w(|x-y|)$ typically follows a Mexican-hat profile, combining short-range excitation with longer-range inhibition. This spatial connectivity pattern enables neural field models to reproduce a rich repertoire of spatiotemporal dynamics including traveling waves, stationary bumps, and spiral patterns observed in cortical tissue [53].

2.1.5 *Large-scale brain network modeling*

The final step in constructing comprehensive brain models involves integrating multiple neural mass or neural field models into large-scale networks that capture the brain's distributed architecture [102, 59, 58]. This approach can be formulated in both continuous and discrete frameworks.

The continuum formulation extends neural field theory to large-scale systems:

$$\begin{aligned} \frac{\partial X(r,t)}{\partial t} &= N(X(r,t)) + \int_{\Omega} w_{\text{local}}(|r-r'|) S(X(r',t))\, dr' \\ &\quad + \int_{\Omega} W_{\text{long-range}}(r,r') S(X(r',t-\tau(r,r')))\, dr' + I(r,t), \end{aligned} \tag{2.8}$$

where the first integral represents local connectivity within a cortical area, and the second integral captures long-range connections between different brain areas with time delays $\tau(r, r') = \tau(|r - r'|)$.

For practical implementation, the discrete formulation represents different brain sub-regions as interconnected nodes:

$$\dot{X}_i(t) = N(X_i(t)) + \sum_j W_{ij} S(X_j(t - \tau_{ij})) + I(t), \tag{2.9}$$

where X_i represents the state vector of region i (which may include multiple variables such as excitatory and inhibitory activities), $N(X_i)$ captures the intrinsic local dynamics of each region, W_{ij} represents the strength of anatomical connections between regions i and j derived from diffusion imaging data, S is a nonlinear transfer function, and τ_{ij} accounts for signal transmission delays proportional to inter-regional distance.

Remark 2.1. *When $N(X_i)$ is taken to be linear, the discrete large-scale network equation reduces to the general neural mass formulation* (2.6)*, now interpreted at the network level rather than for a single population.*

This framework incorporates several crucial neurobiological features. Each brain region maintains its local dynamics while receiving inputs from other regions through the connectome [107, 108, 142, 46]; communication occurs with finite transmission delays reflecting axonal conduction times; and the space-time structure of couplings shapes the emerging network dynamics.

When implementing this framework, researchers typically use established neural mass models such as the Wilson-Cowan equations or more detailed conductance-based models for the local dynamics at each node [142, 46]. The inclusion of transmission delays and noise proves crucial for reproducing realistic resting-state dynamics, enabling the emergence of metastability, multiscale dynamics, and spontaneous switching between functional network configurations observed in empirical data [89].

This integrated approach allows researchers to bridge scales from microscopic neuronal properties to macroscopic brain dynamics, providing a powerful framework for understanding how the brain's anatomical architecture constrains and enables its functional organization across different spatial and temporal scales.

2.2 THE ESTIMATION CHALLENGES IN NEURAL MODELS

The development of neural mass and neural field models provides powerful theoretical frameworks for understanding brain dynamics, but their practical application faces significant estimation challenges. Given that experimental measurements provide only partial observations of the underlying neural states, a fundamental problem in computational neuroscience is estimating the hidden states and parameters that govern these complex dynamical systems.

A central challenge in neural modeling arises from the fact that model parameters are difficult to estimate from experimental data. This is due to several reasons: first, neural mass models employ significant simplifications of biological complexity, collapsing detailed biophysical processes into phenomenological parameters; second, the relationship between model parameters and measurable quantities is often highly nonlinear and non-unique; third, different parameter combinations can produce nearly identical output dynamics, creating identifiability problems. This situation is further complicated by the fact that many parameters, such as synaptic coupling strengths and time constants, cannot be directly measured in vivo and must be inferred indirectly from limited observational data.

From [51]: *A primary challenge is the many degrees of freedom that NFM's (Neural Fields Models which can be generalized to Neural Mass Models (NMM)) have, even with the use of averaged and lumped parameters. This is often due to lack of data and an incomplete understanding of the macroscopic structures and processes. The resulting parameter space is often so large and unconstrained that almost any dynamical behaviour can be constructed. As famously quoted by John Von Neumann: "With four parameters I can fit an elephant, and with five I can make him wiggle his trunk!" This paints an ambiguous physical picture of what is being modelled, as there is often no direct correspondence between model and experimental parameters. For example, it is not currently possible to measure the average synaptic connectivity between two neural populations in the cortex. At best the values used in the literature are order of magnitude estimates that are currently unverifiable.*

Several computational approaches have been developed to solve the identification problem in neural mass and neural field models from experimental data, mainly through *offline* optimization schemes. A classical approach relies on *least-squares estimation* [83], which optimizes parameters by minimizing the discrepancy between simulated model outputs and empirical measurements through iterative gradient-based adjustment. Another classical method is the *grid search*, a straightforward yet conceptually robust approach to parameter estimation that systematically explores the parameter space by evaluating model performance over a discrete set of parameter combinations [125]. Alternatively, *stochastic or deterministic search algorithms*, such as genetic or evolutionary strategies, have been applied to explore the parameter space globally and mitigate local convergence issues [65, 114]. These methods can recover multiple feasible solutions but often become computationally demanding for high-dimensional problems. *Bayesian inference* provides a probabilistic formulation of the estimation problem, yielding posterior distributions of parameters and their uncertainties. Variational Bayesian inversion, as used in Dynamic Causal Modeling (DCM), has been successfully applied to both rate-based and conductance-based neural mass and field models [128, 135]. More recently, hybrid methods have emerged, combining sensitivity analysis and Bayesian inference [151] or using deep-learning-based surrogates to accelerate inversion [155]. These approaches aim to reduce computational cost while maintaining accuracy and uncertainty quantification.

Another important challenge is state estimation. Indeed an important goal is to monitor the internal state of the brain in real time, ranging from tracking seizure dynamics in epilepsy to monitoring anesthesia depth during surgical procedures. Another motivation could be to "enhance" the state estimation by using the model as a "filter" to denoise the measurements and reconstruct the underlying neural activity. One could also imagine increasing the potency of cheap measurement techniques (like EEG) by using the model to reconstruct the underlying neural activity with higher spatial resolution. EEG being a very affordable and widely used technique, this could have important clinical applications.

Beyond monitoring, state estimation provides the foundation for two crucial applications: parameter estimation and feedback control. For parameter estimation, accurate state reconstruction is often a prerequisite for identifying model parameters, as many estimation algorithms rely on knowledge of the complete system state. State observers can provide these necessary estimates, enabling more robust parameter identification. For feedback control applications, such as closed-loop neuro-modulation or brain-computer interfaces, state estimators serve as the "eyes" of the control system, providing the real-time state information needed to compute appropriate control actions [62]. This dual role makes state estimation a cornerstone technology for both understanding and intervening in neural systems.

2.2.1 *Towards online, control-inspired estimation*

Control theory offers a rich set of tools for addressing the limitations of traditional, offline estimation approaches. In particular, nonlinear observer design and adaptive estimation techniques provide a principled framework for the joint state and parameter estimation of a dynamical system, even under partial and noisy observations [124, 121, 31].

One of the key advantages of these online methods is their ability to continuously assimilate new data as it is being measured. Instead of relying on batch computations or full-model inversion, observer-based techniques perform local updates in real time. This makes them particularly well suited for the state estimation problem in neural mass and large-scale models, where nonlinearities, feedback, and noise are pervasive. Moreover, because these methods are grounded in the system's dynamics, they naturally respect temporal continuity and exploit model structure through model-driven updates.

Crucially, online parameter estimation, especially when combined with appropriate excitation, offers a promising route to resolving the parameter estimation problem (neurodegeneracy [20] in computational neuroscience). From the perspective of system identification, different parameter sets may yield indistinguishable outputs unless the system is sufficiently excited. Therefore, to distinguish between competing parameter configurations, one must move beyond passive observation and actively design informative experimental protocols [18, 163]. These may include structured sensory stimuli, targeted electrical inputs, or task-induced fluctuations that ensure each parameter leaves a detectable imprint on the system's output through careful design of excitation protocols.

An additional motivation for online estimation comes from the biological observation that parameters such as synaptic strengths or neuromodulatory gains may themselves vary slowly over time. In contrast to the fast neural state dynamics, these parameters evolve on longer timescales due to plasticity, fatigue, or adaptation. Hence, tracking them requires an estimation framework that is both recursive and temporally adaptive, capable of disentangling multiscale dynamics in real time.

In conclusion, control theory is particularly well suited to address the central challenges of neuroscience because the brain is fundamentally a dynamical, adaptive, and partially observable system operating far from equilibrium [36]. Neural activity emerges from recurrent interactions across multiple spatial and temporal scales, under substantial uncertainty and noise, precisely the setting for which modern estimation and control frameworks were developed

[36]. Concepts such as observability, identification, estimation, robustness, provide a rigorous mathematical language to characterize what can be inferred from neural recordings, what hidden states can be reconstructed, and which perturbations can be filtered. Recent perspectives have increasingly emphasized that many core questions in systems neuroscience can naturally be reframed as state-space estimation and feedback control problems [145, 156, 143, 144]. In particular, control-theoretic approaches offer an important bridge between mechanistic modeling and data assimilation, enabling the integration of biophysical constraints, temporal continuity, and uncertainty quantification into a unified framework.

However, despite this strong conceptual alignment, the methodological toolbox currently available remains insufficient for the scale and complexity of realistic neuroscientific applications [136]. Most observer and adaptive estimation techniques developed in control theory rely on assumptions that are difficult to satisfy in neural systems, including low-dimensional dynamics, strong observability conditions, accurate prior models, or persistent excitation. In practice, neural systems are highly nonlinear, distributed, heterogeneous, and only sparsely observed, while experimental perturbations are often limited by biological and ethical constraints [36, 51]. As a result, robust joint state and parameter estimation in large-scale neural models remains an open challenge. Recent work has highlighted both the promise and the current limitations of adaptive observers and online estimation methods in neuroscience, emphasizing the need for scalable, structure-aware, and uncertainty-aware estimation frameworks capable of handling multiscale neural dynamics and model mismatch [137, 43, 44, 40, 41].

Overall, these considerations point toward the necessity of rigorous and formally grounded methodologies for neural estimation problems. Beyond computational performance alone, future approaches must provide theoretical guarantees on observability, identifiability, stability, and robustness under the severe uncertainty and partial observability that characterize neural systems. Developing such foundations is essential for transforming online estimation and adaptive observer techniques into reliable tools for large-scale neuroscience, and motivates the mathematical perspective adopted in this work.

Part II

STATE ESTIMATION

3

OBSERVABILITY AND OBSERVER DESIGN FOR NEURAL MASS MODELS

The problem of state estimation for neural mass models based on the Wilson-Cowan equations has recently gained increasing attention in the literature, with various approaches proposed for the design of observers for these systems. In the control community, the state estimation analysis of such models has been carried out primarily within the framework of parameter estimation and adaptive observer design [137, 40, 41, 43, 44], while very few studies have addressed the pure state estimation problem. Of those that we are aware of, [48, 120] follow the classical Luenberger observer design with an LMI condition to ensure the convergence. This will be discussed in more detail in the appropriate section.

On the computational side, the state estimation problem for these neural mass models has been investigated in the context of data assimilation, where the objective is to reconstruct the system state from noisy measurements using techniques such as Kalman filtering or variational methods [113, 72, 79, 122, 14] and are usually given without proof of convergence.

The scarcity of such studies is not surprising, given that nearly all existing results require precise knowledge of the system parameters, an assumption that rarely holds in practice. Neural models are well known for being highly uncertain and sensitive to inter-individual variability as well as to slow parameter drifts over longer timescales. We propose nonetheless to study the problem of state estimation for neural mass models on the type (3.1). Designing an observer in this setting may serve as a foundational step towards an adaptive observer, or prove valuable on its own once the parameters have been identified through other means, as discussed in the previous chapter. In this chapter, we recall some known methods in the literature and how they can be applied to neural mass models, especially the Wilson-Cowan model. In the next chapter, we propose a novel LMI condition on the design of a Luenberger observer for this system and give some sufficient conditions and examples.

3.1 MODEL PRESENTATION

We consider the following system

$$\begin{cases} \dot{V} = -DV + WS(V) + Bu, \\ y = CV \quad \text{or} \quad y = C'S(V). \end{cases} \tag{3.1}$$

We recall that here, $V \in \mathbb{R}^n$ represents the state of the system, D is a diagonal matrix with positive entries representing decay rates, W is a synaptic weight matrix, $S(x) = (S_1(x_1), \ldots, S_n(x_n))$, where each S_i is a nonlinear activation function (typically sigmoid or saturation-type), u is an external input, and y is the measured output with C being the measurement matrix.

Remark 3.1. *Note that for this study we can consider systems (3.1) where $D = 0$*

Remark 3.2 (Extension to Lotka-Volterra Systems)**.** *Should the measurement be in term of population, meaning that each component of y is the state of each population, we can extend most of these results for Lotka-Volterra systems, which are of the form*

$$\begin{cases} \dot{x} = D_x(Wx + u_x), \\ y_x = Cx, \end{cases} \tag{3.2}$$

where D_X is the diagonal matrix with diagonal $(x_1, x_2, \cdots, x_n)$. It is well known that the log transformation of this system $z = log(x)$, yields

$$\dot{z} = Wg(z) + u, \tag{3.3}$$

with $g(z) = (e^{z_1}, \cdots, e^{z_n})$.

Methods that do not require the boundedness of S in (3.1) and the matrix D can then be applied for adequate measurement to Lotka-Volterra systems.

Remark 3.3. *Should $D = \lambda I$, through the use of a 2-d sliding mode observer, one can obtain the measurement $C'S(V)$ with $C' := CW$ from the measurement CV without any assumption on the system parameters after a prescribed finite time $T_0 > 0$. Indeed, renaming $z_1 = CV$ and $z_2 = C'S(V)$, one can design a sliding mode observer of the form*

$$\begin{cases} \dot{\hat{z}}_1 = -\lambda z_1 + \hat{z}_2 + Bu + k_1 \operatorname{sign}(z_1 - \hat{z}_1), \\ \dot{\hat{z}}_2 = k_2 \operatorname{sign}(z_1 - \hat{z}_1), \end{cases} \tag{3.4}$$

such that for any time $t \geq T_0$, for k_1, k_2, the estimation error $e = \hat{z} - z$ satisfies $e = 0$ after T_0. This is a direct application of the super-twisting algorithm [117, 26]. See also more details in Appendix 6.6.

3.2 FROM THE AMARI NEURAL FIELD TO A FINITE-DIMENSIONAL MODEL

The observer designs developed in this chapter and in Chapter 4 operate on a finite-dimensional model of the form

$$\begin{cases} \dot{V}(t) = AV(t) + W\,S(V(t)) + Bu(t), \\ y(t) = C\,V(t), \end{cases} \tag{3.5}$$

with $V(t) \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$, $W \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{n_y\times n}$, and $S\colon \mathbb{R}^n \to \mathbb{R}^n$ a bounded saturation. We show how (3.5) arises directly, in continuous time, from the Amari neural field equation together with a linear measurement model representing local field potential (LFP) recordings.

3.2.1 *Continuum model and measurement equation*

We start from the Amari neural field equation [3, 167] on a cortical domain $\Omega \subset \mathbb{R}^d$,

$$\tau \frac{\partial v(r,t)}{\partial t} = -v(r,t) + \int_\Omega w(r,r')\, f\big(v(r',t)\big)\, dr' + I(r,t), \tag{3.6}$$

where $v(r,t) \in \mathbb{R}$ is the mean post-synaptic membrane potential at $r \in \Omega$, $w(r,r')$ is the spatial connectivity kernel, $f\colon \mathbb{R} \to \mathbb{R}$ is the sigmoidal firing-rate of the local population [52], and $I(r,t)$ is an external input current. An electrode placed at $r_n \in \Omega$ measures a spatially weighted average of the membrane potential through an observation kernel m which is generally Gaussian [83, 105],

$$y_n(t) = \int_\Omega m(r_n - r')\, v(r',t)\, dr', \qquad n = 1, \dots, n_y. \tag{3.7}$$

Remark 3.4 (Discrete (Dirac) reduction). *If v is a priori sampled at n collocation nodes $r_1, \dots, r_n \in \Omega$ and the spatial integral in (3.6) is discretised against the measure $d\mu(r') = \sum_{i=1}^n \omega_i\, \delta_{r_i}(r')\, dr'$, then setting $V_i(t) := v(r_i,t)$ and decomposing the kernel as $w(r_i,r_j) = \psi^\top(r_i - r_j)\theta$ gives at once the finite-dimensional system*

$$\tau\, \dot V_i(t) = -V_i(t) + \sum_{j=1}^n \omega_j\, \psi^\top(r_i - r_j)\, \theta\, f(V_j(t)) + I(r_i,t), \tag{3.8}$$

which is exactly (3.5) with $A = -\tau^{-1} I_n$, $W_{ij} = \tau^{-1}\omega_j \psi^\top(r_i - r_j)\theta$, $S(V) = (f(V_1), \dots, f(V_n))^\top$ componentwise, and C the matrix sampling $m * v$ at the sensor positions. This is the same as in (3.1).

3.2.2 *Galerkin reduction on a continuous basis*

For a more faithful continuous reduction, expand the field on n basis functions $\phi_1, \dots, \phi_n \colon \Omega \to \mathbb{R}$ collected in $\phi(r) = (\phi_1(r), \dots, \phi_n(r))^\top$. The ϕ_k are taken as normalised Gaussians $\phi_k(r) \propto \exp(-|r - c_k|^2 / 2\sigma_\phi^2)$ centred on a uniform grid over Ω; the choice of spacing and its relation to the spatial bandwidth of v are discussed in [83].

$$v(r,t) \;\approx\; \phi^\top(r)\, V(t). \tag{3.9}$$

Substituting (3.9) into (3.6), multiplying by $\phi(r)$, and integrating over Ω (Galerkin test) gives the *exact* projected equation

$$\tau\, \Gamma\, \dot V(t) = -\Gamma\, V(t) + \mathcal{F}(V(t)) + b(t), \tag{3.10}$$

where $\Gamma := \int_\Omega \phi(r)\phi^\top(r)\, dr$, $b(t) = \int_\Omega \phi(r)\, I(r,t)\, dr$, and the state-dependent vector $\mathcal{F}(V) \in \mathbb{R}^n$ has components

$$\mathcal{F}_i(V) := \int_\Omega \phi_i(r) \int_\Omega w(r,r')\, f\big(\phi^\top(r')V\big)\, dr'\, dr. \tag{3.11}$$

To obtain a model of the form (3.5) with a *constant* connectivity matrix, we replace $f(\phi^\top(\cdot)V)$ by its $L^2(\Omega)$ projection onto $\mathrm{span}\{\phi_1, \ldots, \phi_n\}$. For a basis that is not orthonormal the L^2 projection of a function $g \in L^2(\Omega)$ onto the span of ϕ is $P_\phi g = \phi^\top \Gamma^{-1} \langle g, \phi \rangle$, so

$$f\big(\phi^\top(r')V\big) \;\approx\; \phi^\top(r')\,\Gamma^{-1}S(V), \qquad S(V) := \int_\Omega \phi(r')\, f\big(\phi^\top(r')V\big)\, dr'. \tag{3.12}$$

Substituting (3.12) into (3.11) yields

$$\tau\,\Gamma\,\dot{V}(t) \approx -\Gamma\, V(t) + \widetilde{W}\,\Gamma^{-1}\, S(V(t)) + b(t), \tag{3.13}$$

where $\widetilde{W} \in \mathbb{R}^{n\times n}$ is the constant bilinear form of the kernel against the basis,

$$\widetilde{W}_{ij} := \int_\Omega \int_\Omega \phi_i(r)\, w(r, r')\, \phi_j(r')\, dr'\, dr. \tag{3.14}$$

Multiplying (3.13) on the left by Γ^{-1}/τ gives (3.5) with

$$A = -\tau^{-1} I_n, \qquad W = \tau^{-1}\Gamma^{-1}\widetilde{W}\,\Gamma^{-1}, \qquad Bu = \tau^{-1}\Gamma^{-1}b(t). \tag{3.15}$$

The nonlinearity $S\colon \mathbb{R}^n \to \mathbb{R}^n$ defined by (3.12) is bounded and inherits the saturation profile of f. Crucially, S is not component-wise: each $S_i(V)$ depends on all components of V through the inner product $\phi^\top(r')V$. The following lemma records its Jacobian bound, used in Chapter 4.

Lemma 3.5 (Sector bound for the Galerkin nonlinearity)**.** *Let $f\colon \mathbb{R} \to \mathbb{R}$ be differentiable with $0 \le f' \le f'_{\max}$ everywhere, and let S be defined by (3.12). Then for every $V \in \mathbb{R}^n$,*

$$0 \;\preceq\; \partial_V S(V) \;=\; \int_\Omega \phi(r')\phi^\top(r')\, f'\big(\phi^\top(r')V\big)\, dr' \;\preceq\; f'_{\max}\,\Gamma, \tag{3.16}$$

and in particular $\|\partial_V S(V)\| \le \gamma$ with the scalar bound $\gamma = f'_{\max}\,\lambda_{\max}(\Gamma)$.

Proof. The integrand $\phi(r')\phi^\top(r')\, f'(\phi^\top(r')V)$ is a positive-semidefinite rank-one matrix for every r', so the integral is $\succeq 0$. Since $f' \le f'_{\max}$, the integral is $\preceq f'_{\max} \int_\Omega \phi\phi^\top = f'_{\max}\,\Gamma$. Taking operator norms yields $\|\partial_V S\| \le f'_{\max}\,\lambda_{\max}(\Gamma) = \gamma$. □

The output equation $y(t) = CV(t)$ follows by injecting (3.9) into (3.7), with $C_{ni} = \int_\Omega m(r_n - r')\phi_i(r')\, dr'$. Comparable basis-function reductions have been used to estimate cortical activity from intracranial recordings [144, 84].

Remark 3.6 (Spectral reduction and the output matrix)**.** *If $\{\phi_k\}$ is the orthonormal eigenbasis of the connectivity operator $\mathcal{W}\colon g \mapsto \int_\Omega w(\cdot, r')g(r')\, dr'$ (the existence of such a basis is guaranteed by the spectral theorem for compact, self-adjoint operators), then $\Gamma = I_n$ and $\widetilde{W} = \mathrm{diag}(\mu_1, \ldots, \mu_n)$, so (3.15) simplifies to $W = \tau^{-1}\mathrm{diag}(\mu_k)$: the reduced connectivity matrix is diagonal, with the eigenvalues of $\mathcal{W}$ as entries. The output matrix becomes for LFP measurements*

$$C_{nk} = \int_\Omega m(r_n - r')\, \phi_k(r')\, dr', \tag{3.17}$$

i.e. the k-th column of C is the LFP kernel m convolved with the k-th eigenfunction of $\mathcal{W}$, sampled at the sensor locations. In general, both the eigenvalues μ_k and the eigenfunctions ϕ_k must be obtained numerically, so neither W nor C is available in closed form.

The Gaussian basis $\phi_k(r) \propto \exp(-|r-c_k|^2/2\sigma_\phi^2)$ sacrifices the diagonal structure of W (at the cost of $\Gamma \neq I_n$ and the Γ^{-1} factors in (3.15)), but yields an output matrix with entries

$$C_{nk} = \int_\Omega m(r_n - r')\,\phi_k(r')\,dr' = \mathcal{N}(r_n - c_k;\, \sigma_m^2 + \sigma_\phi^2)$$

that are explicit, and depend only on the sensor-to-centre distance $\|r_n - c_k\|$. See more details in [83].

3.3 SIMPLE DETECTABILITY CONDITION/ COPY OF THE SYSTEM OBSERVER

We start with the simplest possible observer, which is a direct copy of the system dynamics. Should some structural condition be verified, this simple observer can be shown to converge exponentially.

Proposition 3.7. *Assume $D > 0$ in the sense that each component is strictly positive. Assume S is globally Lipschitz of constant L_S and that $\|W\|L < d_i$ for each $1 \leq i \leq n$. Then, the system (3.1) is contracting and the following observer*

$$\dot{\hat{V}} = -D\hat{V} + WS(\hat{V}) + u, \tag{3.18}$$

ensures exponential convergence of the estimation error to zero.

Proof. Consider the error dynamics $e = \hat{V} - V$. We have

$$\dot{e} = -De + W(S(\hat{V}) - S(V)). \tag{3.19}$$

Using the Lipschitz property of S, we can write

$$\dot{|e|_i} \leq \frac{1}{2}\left(-d_i + \|W\|L\right)|e|_i. \tag{3.20}$$

The condition $\|W\|L < d_i$ ensures exponential decay of the error of each e_i. Therefore, $e(t)$ converges to zero exponentially as $t \to \infty$. □

3.4 CASCADE STRUCTURE OF THE MATRIX W

When the system verify a certain cascade structure, the design of an observer is feasible. One approach is to follow the KKL observer design, which is a powerful theoretical framework for observer design under minimal observability assumptions. In the first subsection we follow the work done in [39], where authors proved that the system (3.1) is backward distinguishable under a cascade structure. The KKL approach requires the numerical inversion of a specific transformation which is generically of dimension $2n+1$. In the next subsection, we show how to design a cascade of sliding-mode observers for the system, which does not require this computation.

3.4.1 *Kazantzis–Kravaris/Luenberger (KKL) approach [39]*

The Kazantzis–Kravaris/Luenberger (KKL) framework provides a general observer design methodology under a minimal observability condition called backward distinguishability [6, 22].

Definition 3.8 (Backward distinguishability, [22])**.** *The system $\dot{x} = f(x,u)$, $y = h(x)$ is backward distinguishable on a set $\mathcal{X} \subset \mathbb{R}^n$ in time $t_d > 0$ if for any $t_0 \geq t_d$ and any pair of solutions $x_a(\cdot)$, $x_b(\cdot)$ with $x_a(t_0) \neq x_b(t_0)$ both in $\mathcal{X}$, there exists $s \in [t_0 - t_d, t_0]$ such that $h(x_a(s)) \neq h(x_b(s))$.*

The KKL observer relies on the existence of an injective mapping $T \colon \mathbb{R}^n \to \mathbb{R}^{n_z}$ (with $n_z \geq 2n+1$ generically) satisfying the PDE

$$\frac{\partial T}{\partial x}(x)\, f(x,u) = \Lambda\, T(x) + \Gamma\, h(x), \tag{3.21}$$

where $\Lambda \in \mathbb{R}^{n_z \times n_z}$ is Hurwitz and $\Gamma \in \mathbb{R}^{n_z \times n_y}$. Defining $z = T(x)$, the dynamics in the new coordinates are linear: $\dot{z} = \Lambda z + \Gamma y$. The observer is then a simple copy

$$\dot{\hat{z}} = \Lambda\, \hat{z} + \Gamma\, y, \qquad \hat{x} = T_L^{-1}(\hat{z}), \tag{3.22}$$

where T_L^{-1} denotes a left-inverse of T. The error $\hat{z} - z$ decays at rate Λ, so $\hat{z}(t) \to z(t)$ exponentially; injectivity of T guarantees $\hat{x}(t) \to x(t)$.

The existence of such an injective T for a generic choice of (Λ, Γ) is guaranteed under backward distinguishability [6, 22]. The main practical difficulty is the computation of T and its left-inverse, which requires solving the PDE (3.21) and typically involves numerical methods [141].

In [39, Remark 3.2], the authors show that the system (3.1) is backward distinguishable under two structural conditions:

1. The pair (W, C) is observable and W is lower triangular.
2. Each activation function S_i is injective.

The lower triangularity of W means that population i receives input only from populations $1, \ldots, i$, giving the system a cascade structure. Combined with observability of (W, C) and the injectivity of S, one can propagate the output information down the cascade and distinguish any two trajectories in backward time.

Under these conditions, the KKL observer (3.22) can be designed. The price to pay is a higher-dimensional observer (usually $n_z \geq 2n+1$) and the need to compute the inverse of the PDE (3.21) numerically. The KKL approach is a powerful theoretical framework for observer design under minimal observability assumptions, but the design of an observer is more often difficult to implement in practice. The numerical computation of the inverse mapping T^{-1} is a non-trivial task, and the resulting observer may be high-dimensional and computationally expensive. Moreover, the structural conditions required for backward distinguishability may not hold in general neural mass models, limiting the applicability of this approach.

3.4.2 *Cascade of sliding-mode observers for chain structure of neural mass models*

In this section, we propose an alternative approach, also relying on a specific cascaded connectivity matrix. Indeed, when the connectivity matrix W has a chain structure, mainly that each population k couples to the next population $k+1$ through a nonzero weight, an explicit cascade of sliding-mode observers [117, 129, 26] can be designed, avoiding the numerical difficulties of the KKL framework. Consider the system in chain form:

$$\begin{cases} \dot{x}_k = -d_k x_k + \sum_{j=1}^{k} w_{kj} S_j(x_j) + w_{k,k+1}\, S_{k+1}(x_{k+1}) + u_k, \quad k = 1,\dots,n-1, \\ \dot{x}_n = -d_n x_n + \sum_{j=1}^{n} w_{nj} S_j(x_j) + u_n, \\ y = x_1, \end{cases} \tag{3.23}$$

where $w_{k,k+1} \neq 0$ for $k = 1,\dots,n-1$. Each population k receives input from populations $1,\dots,k$ (lower triangular part) and from the single next population $k+1$ via the coupling $w_{k,k+1}$. The pair $(W, e_1^\top)$ is observable if and only if all superdiagonal coupling coefficients $w_{k,k+1}$ are nonzero. This form can be obtained by a suitable change of coordinates from any system with an observable pair (W, C) with the pair (W, C) following a lower-triangular structure, as shown in the previous section.

The dissipativity structure ($D > 0$) is not needed for this approach; only the injectivity of each S_k and the boundedness of solutions in a known compact $\mathcal{K}$ are required. Define $y_1 := y = x_1$. For $k = 1,\dots,n-1$, the dynamics of x_k decompose as

$$\dot{x}_k = \underbrace{-d_k x_k + \sum_{j=1}^{k} w_{kj} S_j(x_j) + u_k}_{=:\, f_k(x_k, x_1,\dots,x_{k-1}, t)} + \underbrace{w_{k,k+1}\, S_{k+1}(x_{k+1})}_{=:\, \alpha_k(t)}, \tag{3.24}$$

where f_k is known once populations $1,\dots,k$ have been reconstructed, and α_k is a bounded unknown signal. A super-twisting observer extracts α_k from y_k in finite time. The k-th observer reads:

$$\begin{cases} \dot{\hat{z}}_{k,1} = -d_k \hat{z}_{k,1} + w_{kk}\, S_k(y_k) + \sum_{j=1}^{k-1} w_{kj}\, S_j(y_j) + u_k + \hat{z}_{k,2} + L_k \lceil y_k - \hat{z}_{k,1} \rfloor^{1/2}, \\ \dot{\hat{z}}_{k,2} \in L_k^2 \operatorname{sign}(y_k - \hat{z}_{k,1}), \end{cases} \tag{3.25}$$

where $L_k > 0$ is the observer gain. After the convergence time T_k of the k-th sliding-mode observer, $\hat{z}_{k,2}(t) = \alpha_k(t)$ exactly. The state of population $k+1$ is then recovered as

$$\hat{x}_{k+1} := S_{k+1}^{-1}\left(\frac{\hat{z}_{k,2}}{w_{k,k+1}}\right), \qquad y_{k+1} := \hat{x}_{k+1}, \tag{3.26}$$

and serves as the virtual measurement for the next observer in the chain. Each observer is activated sequentially: observer k is started at time $\bar{T}_{k-1} := \sum_{j=1}^{k-1} T_j$ (with $\bar{T}_0 = 0$), once the previous stage has converged.

Theorem 3.9 (Finite-time cascade reconstruction)**.** *Consider the system (3.23) and suppose:*

1. *Each activation function S_k is injective with $x \mapsto -d_k x + w_{kk} S_k(x)$ at least $\frac{1}{2}$-Hölder continuous.*
2. *The coupling coefficients satisfy $w_{k,k+1} \neq 0$ for $k = 1, \dots, n-1$.*
3. *The solutions of (3.23) are bounded in a known compact set $\mathcal{K} \subset \mathbb{R}^n$.*

Then, for any prescribed convergence times $T_1, \dots, T_{n-1} > 0$, there exist gain thresholds $L_1^*, \dots, L_{n-1}^* > 0$ such that for $L_k > L_k^*$, the cascade observer (3.25)–(3.26) achieves exact finite-time state reconstruction:

$$\hat{x}_{k+1}(t) = x_{k+1}(t), \quad \forall\, t \geq \bar{T}_k := \sum_{j=1}^{k} T_j, \quad k = 1, \dots, n-1. \tag{3.27}$$

In particular, the full state is reconstructed after a total time $T = \sum_{k=1}^{n-1} T_k$.

Proof. The proof proceeds by induction on the cascade index k.

For $k = 1$: the measurement $y_1 = x_1$ is exact. The dynamics read $\dot{x}_1 = f_1(x_1, t) + \alpha_1(t)$ where $f_1(x_1, t) = -d_1 x_1 + w_{11} S_1(x_1) + u_1$ is at least $\frac{1}{2}$-Hölder and $\alpha_1(t) = w_{12}\, S_2(x_2(t))$. Since the solutions are bounded in $\mathcal{K}$, the right-hand side of (3.23) is bounded, so $\dot{x}_2$ is bounded and hence $\dot{\alpha}_1(t) = w_{12}\, S_2'(x_2)\dot{x}_2$ has bounded derivative. By the convergence result for the super-twisting observer with drift [117, 129, 26], there exists L_1^* such that for $L_1 > L_1^*$, $\hat{z}_{1,2}(t) = \alpha_1(t)$ for all $t \geq T_1$. Since S_2 is injective and $w_{12} \neq 0$, the recovery (3.26) yields $\hat{x}_2(t) = x_2(t)$ for $t \geq T_1$.

The inductive step is as follows: suppose $\hat{x}_j(t) = x_j(t)$ for all $j \leq k$ and all $t \geq \bar{T}_{k-1}$, so that $y_k = x_k$ exactly from time $\bar{T}_{k-1}$ onward. Observer block k is activated at $\bar{T}_{k-1}$ with y_k exact. The dynamics of x_k read $\dot{x}_k = f_k(x_k, y_1, \dots, y_{k-1}, t) + \alpha_k(t)$, where f_k is known (all $y_j = x_j$ for $j < k$) and $\frac{1}{2}$-Hölder in x_k, and $\alpha_k = w_{k,k+1}\, S_{k+1}(x_{k+1})$. Since $\dot{x}_{k+1}$ is bounded on $\mathcal{K}$, the derivative $\dot{\alpha}_k = w_{k,k+1}\, S_{k+1}'(x_{k+1})\dot{x}_{k+1}$ is bounded. The same super-twisting argument yields $\hat{z}_{k,2}(t) = \alpha_k(t)$ for $t \geq \bar{T}_{k-1} + T_k = \bar{T}_k$, and injectivity of S_{k+1} gives $\hat{x}_{k+1}(t) = x_{k+1}(t)$ for $t \geq \bar{T}_k$. ☐

Remark 3.10 (Extension to Lotka-Volterra Systems)**.** *Both approaches extend to Lotka-Volterra systems (see remark 3.3) as one can assume $D = 0$ and the functions S_i to be equal to $e^{\cdot}$, which is strictly monotone. For the cascade part, one has only to check that solutions live in a known compact set, which is verified for well tuned parameters.*

3.5 INTERVAL OBSERVER

We consider the system (3.1) with measurement $y = CS(V)$. In this section we recall the classical approach to design an interval observer, i.e. an observer that provides upper and lower bounds on the state. The sufficient condition is the following.

Assumption 3.11. *There exists a matrix $K : \mathbb{R}^{n_y \times n} \to \mathbb{R}^{n \times n_y}$ such that the matrix $(W - KC)$ is Metzler–Hurwitz, meaning that all its off-diagonal components are non-negative and its spectral bound is negative.*

Remark 3.12 (Gain design via linear programming [76, 146]). *By a classical result on cooperative systems [76], a Metzler matrix M is Hurwitz if and only if there exists $v \in \mathbb{R}^n$ with $v > 0$ componentwise such that $Mv < 0$. With this choice both constraints become bilinear in K and v. One could then fix v to a positive vector, and search K. The problem is then*

$$\text{find } K \in \mathbb{R}^{n \times n_y}, v \in (\mathbb{R}_+^*)^n \quad \text{s.t.} \quad (W - KC)_{ij} \geq 0 \ (i \neq j), \quad \sum_j (W - KC)_{ij} v_j \leq -\varepsilon, \quad \forall i, \tag{3.28}$$

One then can either iterate over several positive vectors v or solves the full bilinear problem by alternating between K and v.

Under Assumption 3.11, the following interval observer

$$\begin{cases} \dot{\hat{V}}^+ = -D\hat{V}^+ + WS(\hat{V}^+) + Bu + K(y - CS(\hat{V}^+)), \\ \dot{\hat{V}}^- = -D\hat{V}^- + WS(\hat{V}^-) + Bu + K(y - CS(\hat{V}^-)), \end{cases} \tag{3.29}$$

ensures that for any initial condition such that $\hat{V}_0^- \leq V_0 \leq \hat{V}_0^+$ uniformly, we have $\hat{V}^-(t) \leq V(t) \leq \hat{V}^+(t)$ for all $t \geq 0$ and $\|\hat{V}^+(t) - \hat{V}^-(t)\| \to 0$ as $t \to \infty$. Such initial conditions can always be found, as a system of type (3.1) admits an invariant attractive set due to the boundedness of the non-linear terms and the dissipativity of the linear part. Assuming that the system is initialized in this invariant set, one can then find $\hat{V}_0^-$ and $\hat{V}_0^+$ such that $\hat{V}_0^- \leq V_0 \leq \hat{V}_0^+$. The proof of the convergence of the interval observer relies on the monotonicity properties of the system and the Metzler structure of $(W - KC)$, which ensures that the error dynamics of the upper and lower observers are cooperative and converge to zero. See [76, 94, 126, 139, 67, 66, 146] for more details.

3.5.1 *Numerical Simulation*

We illustrate the interval observer on two coupled E-I pairs, with state $V = (V_{e,1}, V_{e,2}, V_{i,1}, V_{i,2}) \in \mathbb{R}^4$. The connectivity matrix has the biologically motivated block structure

$$W = \begin{pmatrix} W_{ee} & W_{ei} \\ W_{ie} & W_{ii} \end{pmatrix}, \ W_{ee} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}, \ W_{ei} = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}, \ W_{ie} = \begin{pmatrix} 1.5 & 0 \\ 0 & 1.5 \end{pmatrix}, \ W_{ii} = \begin{pmatrix} -4 & 0 \\ 0 & -4 \end{pmatrix},$$

where W_{ei}, W_{ie}, and W_{ii} are diagonal (local connections only) and W_{ee} carries all inter-column excitatory coupling. The decay matrix is $D = I_4$. The measurement is one mixed output per column,

$$y_j = S_e(V_{e,j}) - \tfrac{1}{2} S_i(V_{i,j}), \quad j = 1, 2, \qquad C = \begin{pmatrix} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -\frac{1}{2} \end{pmatrix},$$

with sigmoid activation $S_k(v) = (1 + e^{-v})^{-1}$. The parameters here have been chosen for theoretical illustration purposes, not to fit any specific biological data.

The gain $K = (K_e^\top, K_i^\top)^\top$ with

$$K_e = \begin{pmatrix} 5 & 0.5 \\ 0.5 & 5 \end{pmatrix}, \qquad K_i = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

gives

$$W - KC = \begin{pmatrix} -2 & 0.5 & 0.5 & 0.25 \\ 0.5 & -2 & 0.25 & 0.5 \\ 2.5 & 0 & -4.5 & 0 \\ 0 & 2.5 & 0 & -4.5 \end{pmatrix}. \tag{3.30}$$

All off-diagonal entries are non-negative (Metzler) and all row sums (-0.75, -0.75, -2, -2) are negative (Hurwitz). Figure 3.1 shows the results over $t \in [0, 5]$ with true initial condition $V_0 = (0.3, -0.2, 0.1, -0.1)^\top$ and bounding initial conditions $\hat{V}_0^+ = (1, 1, 1, 1)^\top$, $\hat{V}_0^- = (-1, -1, -1, -1)^\top$. The simulations were carried out using a Runge–Kutta RK45 method with time step $dt = 10^{-3}$

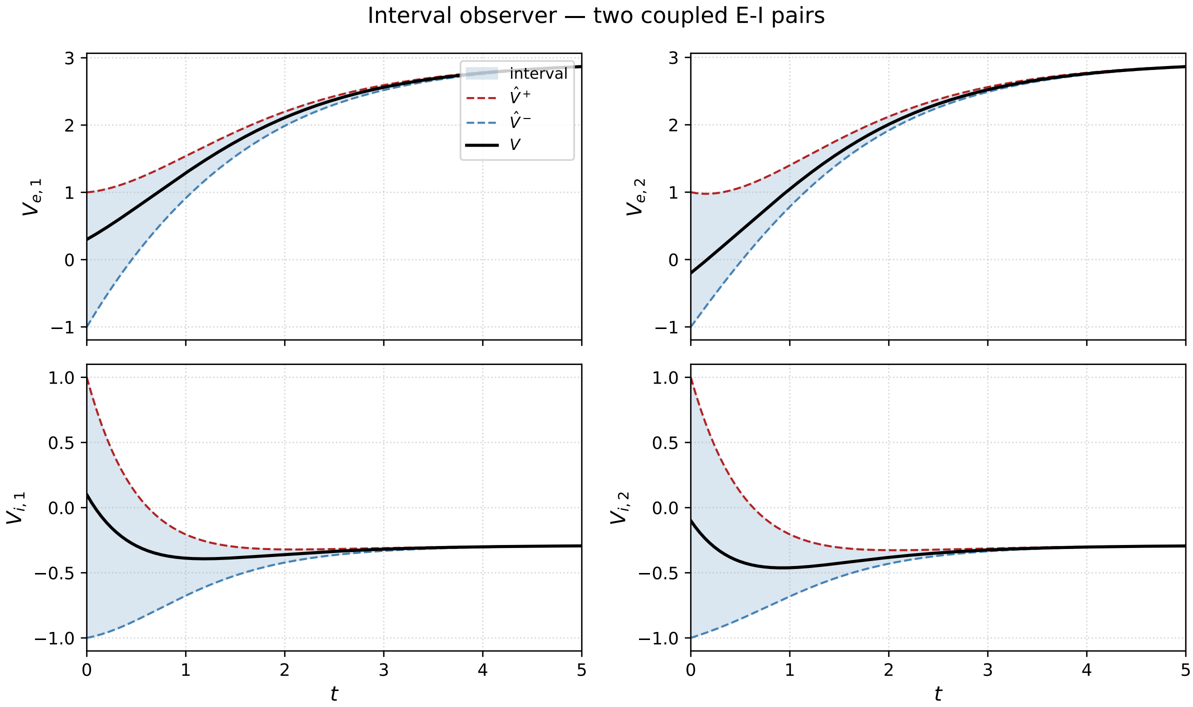


Figure 3.1: Interval observer on two coupled E-I (Excitation-Inhibition) population. Each panel shows one component: true trajectory $V_k(t)$ (black), upper bound $\hat{V}_k^+(t)$ (red, dashed), lower bound $\hat{V}_k^-(t)$ (blue, dashed). The bounds bracket the true state at all times and converge to it asymptotically.

Remark 3.13. *The interval observer and its convergence proof require neither $D \succ 0$ nor boundedness of S. The argument relies only on the Metzler structure of $W - KC$ and the monotonicity of S. Setting $D = 0$ and replacing S by any strictly monotone function leaves the proof intact. In particular, the construction applies to Lotka-Volterra systems as described in Remark 3.2, where $S_i(v) = e^v$ is strictly monotone, solutions are bounded in a known compact set for well-tuned parameters, and the log-transformed system has the same structure as (3.1) with $D = 0$.*

3.6 GENERAL LUR'E SYSTEM LMI-BASED OBSERVER DESIGN [12, 170, 90]

Linear Matrix Inequality (LMI) approaches provide a powerful yet simple framework for observer design. Given the system and a candidate observer depending on tunable matrices, one writes the error dynamics and seeks a quadratic Lyapunov function $V(e) = e^\top Pe$ to ensure convergence. The condition such that the error dynamic is asymptotically stable becomes a condition on the matrix P of the Lyapunov function and the matrices of the observer. The condition for the dynamic of the error to be contracting then becomes a linear matrix inequality [35]. Since LMI feasibility can be checked by semidefinite programming, the observer design reduces to a convex optimization problem. Systems of the form (3.1) can be seen as a specific case of a more general class of systems known as Lur'e systems, which consist of a linear time-invariant part interconnected with a static nonlinearity. The observer design for such systems has been extensively studied in the literature [35, 90, 12, 170, 75], and LMI-based methods have been developed to provide systematic conditions for the existence of observers that ensure exponential convergence of the estimation error.

This section is inspired by the work of [90, 12, 170, 75] and the review article [23]. Recalling their work, we consider a general class of systems of the form

$$\dot{x} = Ax + G\phi(\zeta, y, t) + Bu, \quad y = Cx, \quad \zeta = Mx, \tag{3.31}$$

where $x \in \mathbb{R}^n$, $y \in \mathbb{R}^{n_y}$, $\zeta \in \mathbb{R}^{n_\zeta}$, and (A, G, M, C, B) are matrices of appropriate dimensions. The function $\phi\colon \mathbb{R}^{n_\zeta} \times \mathbb{R}^{n_y} \times \mathbb{R} \to \mathbb{R}^{n_\zeta}$ is smooth and acts component-wise.

The observers considered in [12, 75] take the form

$$\begin{cases} \dot{\hat{x}} = A\hat{x} + K(y - C\hat{x}) + G\phi(\hat{\zeta}, y, t) + Bu, \\ \hat{\zeta} = M\hat{x} + E(y - C\hat{x}), \end{cases} \tag{3.32}$$

where $K \in \mathbb{R}^{n \times n_y}$ is the observer gain and $E \in \mathbb{R}^{n_\zeta \times n_y}$ modifies the argument at which the nonlinearity is evaluated. When $E = 0$, one recovers a classical Luenberger-type observer; the matrix E provides an additional degree of freedom by injecting the output correction into the nonlinearity argument.

The estimation error $e = \hat{x} - x$ satisfies

$$\dot{e} = (A - KC)\, e + G\, \delta, \tag{3.33}$$

where $\delta = \phi(\hat{\zeta}, y, t) - \phi(\zeta, y, t)$ and $\hat{\zeta} - \zeta = (M - EC)\, e$.

A sufficient condition for the observer to be uniformly exponentially convergent is to ensure that the dynamic of the error is contracting. A sufficient condition for this is the existence of $P > 0$ and $q > 0$ such that

$$\mathrm{He}\left\{ P \left[A - KC + G \frac{\partial \phi}{\partial \zeta}(\zeta, y, t)\, (M - EC) \right] \right\} \leq -qP, \tag{3.34}$$

for all (ζ, y, t). This is an infinite-dimensional condition, since it must hold pointwise over all values of (ζ, y, t). Imposing structural assumptions on ϕ reduces it to a finite-dimensional LMI.

3.6.1 *Globally Lipschitz nonlinearities*

Assume that ϕ satisfies the global Lipschitz condition

$$\left| G \frac{\partial \phi}{\partial \zeta}(\zeta, y, t)\, M \right| \leq \bar{b}, \quad \forall\, (\zeta, y, t), \tag{3.35}$$

for some $\bar{b} > 0$. Then, taking $E = 0$, a sufficient condition for (3.34) is the existence of $P > 0$ and $R \in \mathbb{R}^{n_y \times n}$ such that

$$\begin{pmatrix} A^\top P + PA - R^\top C - C^\top R + I & P \\ P & -\frac{1}{\bar{b}}\, I \end{pmatrix} \leq 0, \tag{3.36}$$

with the observer gain selected as $K = P^{-1} R^\top$ [154, 140]. This is a standard LMI in the decision variables (P, R) and can be solved by semidefinite programming.

Remark 3.14. *The simple contraction condition of the previous section is a particular case. For the Wilson-Cowan system (3.1) with $A = -D$, $G = W$, $M = I$, the Lipschitz bound is $\bar{b} = \|W\|\,\gamma$ where $\gamma = \|S'\|_\infty$. Taking $K = 0$ ($R = 0$) and $P = I$ in (3.36), the Schur complement yields $-2D + (1 + \bar{b})I \leq 0$, recovering a contraction condition of the form $d_i > (\|W\|\gamma + 1)/2$. The condition $\|W\|_\infty \gamma < d_i$ obtained earlier corresponds to the same phenomenon with a different choice of Lyapunov metric.*

3.6.2 *Monotone nonlinearities*

When $\phi\colon \mathbb{R}^{n_\zeta} \to \mathbb{R}^{n_\zeta}$ is monotone, namely

$$(\zeta_a - \zeta_b)^\top [\phi(\zeta_a, y, t) - \phi(\zeta_b, y, t)] \geq 0, \quad \forall\, (\zeta_a, \zeta_b, y, t), \tag{3.37}$$

the increment $\delta = \phi(\hat{\zeta}, y, t) - \phi(\zeta, y, t)$ satisfies the single quadratic constraint $\delta^\top (\hat{\zeta} - \zeta) \geq 0$.

We recall the following result from [35], known in the control community as the S-procedure.

Lemma 3.15 (S-procedure, [35])**.** *Let $f_0, f_1, \dots, f_p\colon \mathbb{R}^N \to \mathbb{R}$ be quadratic functions. Suppose one wishes to show that $f_0(x) < 0$ for all $x \neq 0$ satisfying $f_i(x) \leq 0$, $i = 1, \dots, p$. A sufficient condition is the existence of multipliers $\tau_1, \dots, \tau_p \geq 0$ such that*

$$f_0(x) - \sum_{i=1}^{p} \tau_i f_i(x) < 0, \quad \forall\, x \neq 0.$$

When $p = 1$, the condition is also necessary (lossless S-procedure).

Since $\hat{\zeta} - \zeta = (M - EC)e$, this is the constraint $\delta^\top (M - EC)e \geq 0$. Applying the S-procedure (Lemma 3.15) with a scalar multiplier $\lambda > 0$ and combining with the Lyapunov decrease $\dot{V} \leq -q\|e\|^2$, the condition (3.34) is satisfied if there exist $P > 0$, $q > 0$, $\lambda > 0$, and matrices K, E such that

$$\begin{pmatrix} (A - KC)^\top P + P(A - KC) + qI & PG + \lambda\, (M - EC)^\top \\ G^\top P + \lambda\, (M - EC) & 0 \end{pmatrix} \leq 0. \tag{3.38}$$

The zero $(2,2)$ block forces $PG + \lambda(M - EC)^\top = 0$ exactly, so λ is fixed rather than free: one chooses $\lambda > 0$ and solves for (P, K, E) subject to (3.38).

Remark 3.16 (Component-wise monotonicity)**.** *When each ϕ_i depends only on ζ_i and is nondecreasing in that argument, the single constraint $\delta^\top(M-EC)e \geq 0$ decomposes into n_ζ independent scalar constraints $\delta_i[(M-EC)e]_i \geq 0$. Applying the S-procedure with a free diagonal multiplier $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_{n_\zeta}) > 0$ (one λ_i per constraint) replaces λ by Λ in (3.38), recovering the observer LMI of [12, 75]. The diagonal structure of Λ makes the condition less conservative than fixing a scalar λ.*

3.6.3 *Global sector-bounded nonlinearities*

When the Jacobian of ϕ with respect to ζ is symmetric and satisfies the matrix bound

$$0 \preceq \frac{\partial\phi}{\partial\zeta}(\zeta, y, t) \preceq \gamma I, \quad \forall(\zeta, y, t), \tag{3.39}$$

the increment $\delta = \phi(\hat{\zeta}, y, t) - \phi(\zeta, y, t)$ satisfies, by the mean-value theorem, $\delta = \bar{J}(\hat{\zeta} - \zeta)$ where $\bar{J} = \int_0^1 \frac{\partial\phi}{\partial\zeta}(\zeta + t(\hat{\zeta} - \zeta), y, t)\, dt$ satisfies $0 \preceq \bar{J} \preceq \gamma I$, so

$$\delta^\top\big(\delta - \gamma\,(M-EC)\,e\big) \leq 0, \tag{3.40}$$

since $\delta^\top\delta = e^\top(M-EC)^\top\bar{J}^2(M-EC)e \leq \gamma\, e^\top(M-EC)^\top\bar{J}(M-EC)e = \gamma\,\delta^\top(M-EC)e$. Here, we assumed for simplicity that $\bar{J}$ is symmetric.

The bound (3.39) does *not* require ϕ to act component-wise; it holds whenever the Jacobian is bounded between zero and γI in the PSD ordering.

Applying the S-procedure (Lemma 3.15) with the single constraint (3.40) and a scalar multiplier $\lambda > 0$, one obtains the following.

Proposition 3.17. *If there exist $P > 0$, $q > 0$, $\lambda > 0$, and matrices K, E such that*

$$\begin{pmatrix} (A-KC)^\top P + P(A-KC) + qI & PG + \lambda\gamma\,(M-EC)^\top \\ G^\top P + \lambda\gamma\,(M-EC) & -2\lambda I \end{pmatrix} \leq 0, \tag{3.41}$$

then the observer (3.32) achieves global exponential convergence of the estimation error to zero.

Proof. Consider $\mathcal{V}(e) = e^\top Pe$. Along (3.33),

$$\dot{\mathcal{V}} = e^\top \mathrm{He}\{P(A-KC)\}\, e + 2\, e^\top PG\,\delta.$$

Applying the S-procedure with the single constraint (3.40) and multiplier $\lambda > 0$, a sufficient condition for $\dot{\mathcal{V}} \leq -q\|e\|^2$ is

$$\dot{\mathcal{V}} - 2\lambda\,\delta^\top\big(\delta - \gamma(M-EC)e\big) \leq -q\|e\|^2 \quad \forall(e, \delta).$$

Expanding: the term $-2\lambda\|\delta\|^2$ produces the $-2\lambda I$ block in the $(2,2)$ position, and the cross-term $2\lambda\gamma\,\delta^\top(M-EC)e$ produces $\lambda\gamma(M-EC)^\top$ in the off-diagonal. The resulting matrix inequality is (3.41). □

The following subsection shows that when ϕ *does* act component-wise, the single constraint (3.40) splits into n_ζ independent ones, recovering a free diagonal multiplier Λ and a less conservative LMI.

3.6.4 *Component-wise sector-bounded nonlinearities*

When each ϕ_i depends only on ζ_i and is slope-restricted in the sector $[0, \gamma]$:

$$0 \leq \frac{\phi_i(\zeta_a, y, t) - \phi_i(\zeta_b, y, t)}{\zeta_a - \zeta_b} \leq \gamma, \quad \forall \zeta_a \neq \zeta_b, \tag{3.42}$$

both bounds can be exploited independently for each channel. Writing $\delta_i = \phi_i(\hat{\zeta}_i) - \phi_i(\zeta_i)$ and $\hat{\zeta} - \zeta = (M - EC)\, e$, the component-wise sector condition yields

$$\delta_i\big(\delta_i - \gamma\, [(M - EC)\, e]_i\big) \leq 0, \quad i = 1, \ldots, n_\zeta. \tag{3.43}$$

Applying the same approach as in the monotone case, but now with one multiplier $\lambda_i > 0$ per sector constraint (3.43), one obtains the following.

Proposition 3.18 ([12, 75, 170]). *If there exist $P > 0$, $q > 0$, a diagonal $\Lambda > 0$, and matrices K, E such that*

$$\begin{pmatrix} (A - KC)^\top P + P(A - KC) + qI & PG + \gamma\, (M - EC)^\top \Lambda \\ G^\top P + \gamma\, \Lambda (M - EC) & -2\Lambda \end{pmatrix} \leq 0, \tag{3.44}$$

then the observer (3.32) achieves global exponential convergence of the estimation error to zero.

Proof. Consider $V(e) = e^\top Pe$. Along (3.33),

$$\dot{V} = e^\top \mathrm{He}\{P(A - KC)\}\, e + 2\, e^\top PG\, \delta. \tag{3.45}$$

We wish to show $\dot{V} \leq -q\|e\|^2$ under the n_ζ sector constraints (3.43). Applying the S-procedure with one multiplier $\lambda_i > 0$ per constraint, a sufficient condition is

$$\dot{V} - \sum_{i=1}^{n_\zeta} 2\lambda_i\, \delta_i\big(\delta_i - \gamma\, [(M - EC)\, e]_i\big) \leq -q\|e\|^2 \quad \forall\, (e, \delta) \neq 0. \tag{3.46}$$

Expanding the left-hand side, the cross-terms $2\lambda_i \gamma\, \delta_i [(M - EC)e]_i$ produce the off-diagonal blocks $\gamma(M - EC)^\top \Lambda$ in the matrix, while the terms $-2\lambda_i \delta_i^2$ produce the -2Λ block in the $(2, 2)$ position. The full expression rewrites as

$$\begin{pmatrix} e \\ \delta \end{pmatrix}^\top \begin{pmatrix} \mathrm{He}\{P(A - KC)\} + qI & PG + \gamma\, (M - EC)^\top \Lambda \\ G^\top P + \gamma\, \Lambda (M - EC) & -2\Lambda \end{pmatrix} \begin{pmatrix} e \\ \delta \end{pmatrix} \leq 0, \tag{3.47}$$

which is guaranteed by negative semi-definiteness of the matrix, i.e. (3.44). □

The crucial difference between this result and the globally Lipschitz LMI (3.36) lies in how the nonlinearity increment $\delta = \phi(\hat{\zeta}) - \phi(\zeta)$ is bounded. The Lipschitz approach treats ϕ as a black box and uses a single scalar bound $\|\delta\| \leq \bar{b}\, \|e\|$ which is a norm-level inequality that couples all components together and discards the fact that ϕ acts component-wise. By contrast, the sector approach exploits the component-wise structure. Each δ_i is individually bounded in terms of the corresponding $(\hat{\zeta}_i - \zeta_i)$, and the diagonal multiplier $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_{n_\zeta})$ assigns an independent weight to each channel. This channel-by-channel treatment provides n_ζ

independent degrees of freedom instead of one, yielding a less conservative feasibility region. The -2Λ block in the $(2,2)$ position, originating from the upper slope bound γ, provides negative definiteness in the δ direction, which is absent in the monotone LMI (3.38) where only the lower bound $\phi_i' \geq 0$ is available.

With the change of variable $Y = PK$, the terms $A^\top P + PA - C^\top Y^\top - YC$ in the $(1,1)$ block become linear in (P, Y). When $E = 0$, the full LMI (3.44) is linear in the decision variables (P, Y, Λ) and can be solved directly. When $E \neq 0$, the product EC in $(M - EC)^\top \Lambda$ introduces a bilinearity with Λ; this can be handled by fixing Λ and alternating, or by an additional change of variable as discussed in [75, 90].

Remark 3.19 (LMI versus BMI). *The condition (3.44) is a matrix inequality in the unknowns (P, K, E, Λ). A matrix inequality $\mathcal{M}(x) \leq 0$ is a genuine LMI (solvable by semidefinite programming in polynomial time) if and only if $\mathcal{M}$ depends affinely on the decision variables. Products of two unknowns yield a bilinear matrix inequality (BMI), which is non-convex in general.*

In (3.44), the product PK is linearized by the substitution $Y = PK$ (valid since $P > 0$). However, the off-diagonal block contains $(M - EC)^\top \Lambda$, which involves the product $E^\top \Lambda$ which is bilinear in (E, Λ). Consequently:

1. *When $E = 0$: the condition (3.44) is a genuine LMI in (P, Y, Λ).*
2. *When $E \neq 0$ and Λ is fixed: the condition is an LMI in (P, Y, E).*
3. *When both E and Λ are free: the condition is a BMI.*

When Λ must be fixed, a natural choice is $\Lambda = \alpha I$ for some $\alpha > 0$, which assigns equal weight to all sector constraints. One can then perform a line search over α: for each candidate α, solve the resulting LMI in (P, Y, E) and check feasibility. Alternatively, one can alternate between fixing Λ and solving for (P, Y, E), then fixing (P, E) and solving for (Y, Λ), iterating until convergence. Neither strategy is guaranteed to find a global optimum, but both are effective in practice.

The same distinction applies to all subsequent LMIs in this chapter: when E (or any nonlinearity-argument modifier) is absent, the S-procedure multiplier Λ appears affinely alongside (P, K) and the condition is a genuine LMI.

An equivalent and perhaps more transparent reformulation, due to Zemouche and Boutayeb [170], is obtained for the case $M = I$, $E = 0$ by observing that the sector condition directly yields a polytopic structure on the error dynamics. Since each ϕ_i has slope in $[0, \gamma]$, the increment decomposes as $\delta_i = \phi_i(\hat{\zeta}_i) - \phi_i(\zeta_i) = \Delta_i \cdot (\hat{\zeta}_i - \zeta_i)$ with $\Delta_i \in [0, \gamma]$. When $M = I$ and $E = 0$, we have $\hat{\zeta} - \zeta = e$, so $\delta = \Delta e$ with $\Delta = \mathrm{diag}(\Delta_1, \ldots, \Delta_n) \in [0, \gamma I]$. Substituting into the error dynamics (3.33) yields

$$\dot{e} = (A - KC + G\Delta)\, e, \tag{3.48}$$

a linear time-varying system where $\Delta(t)$ ranges over the hypercube $[0, \gamma]^n$. Since this set is a convex polytope, quadratic stability for all $\Delta \in [0, \gamma I]$ is equivalent to checking the 2^n vertices $\Delta_v \in \{0, \gamma\}^n$:

$$\mathrm{He}\{P(A - KC + G\Delta_v)\} + qI \leq 0, \quad \forall\, \Delta_v \in \{0, \gamma\}^n. \tag{3.49}$$

This vertex approach and the S-procedure multiplier approach (3.44) are equivalent formulations of the same condition [12, 170].

3.6.5 *Application to Wilson-Cowan equations in the general case*

The Wilson-Cowan system (3.1) is of the form (3.31) with

$$A = -D, \quad G = W, \quad M = I, \quad \phi = S, \quad C = C. \tag{3.50}$$

Since the sigmoid S is monotone and slope-restricted in $[0, \gamma]$ with $\gamma = \|S'\|_\infty$, the sector-bounded LMI (3.44) applies. The observer (3.32) reads

$$\begin{cases} \dot{\hat{V}} = -D\hat{V} + K(y - C\hat{V}) + WS\big(\hat{V} + E(y - C\hat{V})\big) + u, \\ \hat{\zeta} = \hat{V} + E(y - C\hat{V}), \end{cases} \tag{3.51}$$

and the LMI (3.44) becomes

$$\begin{pmatrix} (-D-KC)^\top P + P(-D-KC) + qI & PW + \gamma\,(I - EC)^\top \lambda I \\ W^\top P + \gamma\,\lambda(I - EC) & -2\lambda I \end{pmatrix} \leq 0. \tag{3.52}$$

This is a direct application of the general framework. The gain K acts on the linear output error $(y - C\hat{V})$, injecting correction into the linear part of the dynamics. The matrix E provides an additional degree of freedom by modifying the nonlinearity argument. When P is a full matrix, the LMI has (P, K, E, Λ) as decision variables, potentially coupled through bilinear terms.

The direct application however does not usually yield feasibility in the more general case $\dot{V} = AV + WS(V) + Bu$. Chong et al. [48] apply the circle criterion LMI framework to a specific single cortical column neural mass model and find that the standard conditions are infeasible for that particular system. They generalise the LMI by introducing an additional multiplier matrix to recover feasibility, and also account for input uncertainty and measurement noise. In [120], another LMI is derived for general Lur'e systems with sector-bounded nonlinearities, and applied to a specific neural mass model. The resulting LMI is more complex than (3.52) and involves additional decision variables and multipliers. The authors show that it can perform better in certain conditions than the standard LMI (3.52) or the one proposed in [48]. In the next chapter we will propose a novel approach to design an observer for these neural mass models (3.31) that exploits the possibility to reconstruct through a finite-dimensional exact differentiator a second *virtual output* which would allow us to further relax the LMI conditions and enlarge the feasibility region.

4

OBSERVER DESIGN FOR LUR'E SYSTEMS VIA INJECTION OF RECONSTRUCTED OUTPUT

We consider the problem of state estimation for neural mass models of the form $\dot{V} = -DV + WS(V) + Bu$, $y = CV$, viewed as Lur'e systems with sector-bounded nonlinearities. In the classical LMI-based observer, a single gain K injects the linear output error $y - C\hat{V}$ into the dynamics; the connectivity matrix W appears as a fixed parameter in the resulting matrix inequality and cannot be shaped by the design. We show that, when the dissipation is uniform ($D = \lambda I$) or the measurement matrix C selects individual populations, n_y parallel super-twisting observers reconstruct the nonlinear signal $y_2 = CWS(V)$ exactly in prescribed finite time. Using this additional measurement, we introduce a nonlinear injection gain K' that replaces the effective connectivity by $(I - K'C)W$ in the error dynamics while leaving the dissipation $-D$ unchanged. The combined observer, with design variables (K, K'), leads to an LMI whose feasibility region strictly contains that of the standard formulation (recovered when $K' = 0$). Numerical simulation are showcased and the structural advantages of the nonlinear injection approach are discussed. We further improve on this result on the general case $\dot{V} = AV + WS(V) + Bu$. We finally test the observer to reconstruct the state of neural fields from a finite number of measurement channels.

To the best of our knowledge, combining exact finite-time reconstruction of a nonlinear output with a nonlinear injection observer design is novel in the literature.

4.1 INTRODUCTION

In the previous chapter, the neural mass model (3.1) was cast as a Lur'e system with $A = -D$, $G = W$, $M = I$, and the sector-bounded LMI (3.52) was derived as a direct application of the general framework of [12, 170, 75]. In that formulation, the observer gain K enters the $(1, 1)$ block of the LMI through $\mathrm{He}\{P(-D - KC)\}$. Increasing K accelerates the linear damping but simultaneously modifies the balance between dissipation and nonlinear coupling. The nonlinearity $WS(V)$ is bounded only via the sector condition, and the gain has no direct action on the coupling matrix W itself.

The starting point is that, under a uniform decay rate $D = \lambda I$ (or, more generally, when each row of C is supported on indices sharing a common d_i), n_y parallel super-twisting observers reconstruct the signal $CWS(V(t))$ exactly in prescribed finite time T_0. After T_0, two independent measurements are available: $y = CV$ (linear in the state) and $y_2 = CWS(V)$

(nonlinear). A gain K' can then correct the predicted nonlinear output $CWS(\hat{V})$ against the measured y_2, replacing W by $(I - K'C)W$ in the error dynamics while leaving $-D$ untouched. This structural decoupling is absent from the standard formulation.

The chapter is organized as follows. Section 4.2 formalizes the finite-time reconstruction of y_2. Section 4.3 derives the nonlinear injection observer and its LMI conditions. Section 4.4 combines linear and nonlinear injection into a unified observer with design matrices (K, K', E). Section 4.5 discusses the structural advantages over the naive application of the general Lur'e framework.

We recall the system under consideration:

$$\begin{cases} \dot{V} = -DV + WS(V) + Bu, \\ y = CV, \end{cases} \tag{4.1}$$

where $V \in \mathbb{R}^n$, $D = \operatorname{diag}(d_1, \ldots, d_n)$ with $d_i > 0$, $W \in \mathbb{R}^{n\times n}$ is the connectivity matrix, $S\colon \mathbb{R}^n \to \mathbb{R}^n$ acts component-wise with each S_i monotone and slope-restricted in $[0, \gamma_i]$ ($\gamma = \max_i \gamma_i$), $B \in \mathbb{R}^{n\times n_u}$, and $C \in \mathbb{R}^{n_y \times n}$ is the measurement matrix. We assume that solutions are bounded in a known compact set.

4.2 SLIDING-MODE RECONSTRUCTION OF THE NONLINEAR OUTPUT

In this section, we show that the signal $y_2 = CWS(V)$ can be reconstructed exactly in finite time by n_y parallel super-twisting observers which act as exact differentiators. For this construction, we restrict ourselves to a uniform decay rate

$$D = \lambda\, I_n, \qquad \lambda > 0. \tag{4.2}$$

Under this condition, applying C to (4.1) and setting $z_1 := y$, $z_2 := CWS(V)$ gives

$$\dot{z}_1 = -\lambda\, z_1 + z_2 + CBu. \tag{4.3}$$

For a general diagonal $D = \operatorname{diag}(d_1, \ldots, d_n)$, the term CDV cannot be written as a function of y alone unless $\ker C$ is D-invariant, i.e. each row of C is supported on indices sharing a common d_i. The uniform case (4.2) is the only one that handles arbitrary C; the case where each row of C selects an individual population is addressed in the remark below.

Here z_1 is measured, and z_2 is an unknown signal, bounded with bounded derivative (by forward invariance of solutions in a compact set). For each measured component j, the scalar subsystem

$$\dot{z}_{1,j} = -\lambda\, z_{1,j} + z_{2,j} + [CBu]_j \tag{4.4}$$

has the standard form amenable to super-twisting reconstruction. The observer

$$\begin{cases} \dot{\hat{z}}_{1,j} = -\lambda\, \hat{z}_{1,j} + \hat{z}_{2,j} + [CBu]_j + L_j \lceil z_{1,j} - \hat{z}_{1,j} \rfloor^{1/2}, \\ \dot{\hat{z}}_{2,j} \in L_j^2 \operatorname{sign}(z_{1,j} - \hat{z}_{1,j}), \end{cases} \tag{4.5}$$

achieves finite-time convergence for L_j above a computable threshold L_j^* that depends on $\|z_{2,j}\|_\infty$ and $\|\dot{z}_{2,j}\|_\infty$ [117, 129, 26]. The drift $-\lambda\,\hat{z}_{1,j} + [CBu]_j$ is $\frac{1}{2}$-Hölder continuous in $\hat{z}_{1,j}$, satisfying the regularity condition of the super-twisting convergence result with drift. Running n_y such observers in parallel reconstructs the full vector

$$y_2(t) := CWS(V(t)) \tag{4.6}$$

exactly for all $t \geq T_0 := \max_j T_{0,j}$.

Proposition 4.1 (Finite-time reconstruction of y_2)**.** *Consider the system (4.1) with $D = \lambda I_n$, C a measurement matrix and solutions bounded in a known compact set. For any prescribed $T_0 > 0$, there exist gains $L_1, \dots, L_{n_y}$ such that the n_y parallel super-twisting observers (4.5) satisfy $\hat{z}_2(t) = CWS(V(t))$ for all $t \geq T_0$.*

Remark 4.2 (Population-based measurements)**.** *When the measurement matrix C selects individual populations, i.e. each row of C is a standard basis vector $e_{j_k}^\top$, the activation function S acts component-wise so that $S(y) = S(CV) = CS(V)$. In this case the nonlinear output $CS(V)$ is directly available from the measurement y and the sliding-mode reconstruction above is not needed to obtain a second virtual measurement. One still can introduce this finite reconstruction to obtain an additional third measurement of the form $CWS(V)$, which may further relax the LMI conditions, but this is beyond the scope.*

On the other hand, the uniform-decay assumption (4.2) can be dropped and the observer of Section 4.3 applies for general diagonal D.

We assume hereafter that $t \geq T_0$, so that both $y = CV$ and $y_2 = CWS(V)$ are available.

4.3 OBSERVER WITH NONLINEAR OUTPUT INJECTION

Consider the observer

$$\dot{\hat{V}} = -D\hat{V} + WS(\hat{V}) - K'(CWS(\hat{V}) - y_2) + Bu, \tag{4.7}$$

where $K' \in \mathbb{R}^{n \times n_y}$ is the design gain. The estimation error $e = \hat{V} - V$ satisfies

$$\dot{e} = -De + (W - K'CW)\,\delta, \tag{4.8}$$

where $\delta = S(\hat{V}) - S(V)$. By the mean value theorem, $\delta_i = \Delta_i\, e_i$ with $\Delta_i \in [0, \gamma]$, so that

$$\dot{e} = \big(-D + (I - K'C)W\,\Delta_e\big)\,e, \qquad \Delta_e = \operatorname{diag}(\Delta_1, \dots, \Delta_n) \in [0, \gamma I]. \tag{4.9}$$

The gain K' replaces the effective connectivity W by $(I - K'C)W$ without affecting the dissipation $-D$.

Quadratic stability for all vertices $\Delta_v \in \{0, \gamma\}^n$ reads

$$\operatorname{He}\{P(-D + (W - K'CW)\Delta_v)\} + qI \leq 0, \quad \forall\, \Delta_v \in \{0, \gamma\}^n. \tag{4.10}$$

Following the same approach as in Section 3.6, one can incorporate heterogeneous slopes $\gamma_i = \|S_i'\|_\infty$ through a diagonal multiplier $\Lambda > 0$ and the matrix $\Gamma = \mathrm{diag}(\gamma_1, \ldots, \gamma_n)$, leading to

$$\begin{pmatrix} -(PD+DP)+qI & P(W-K'CW)+\Gamma\,\Lambda \\ (W-K'CW)^\top P+\Lambda\,\Gamma & -2\Lambda \end{pmatrix} \leq 0. \tag{4.11}$$

Proposition 4.3. *If there exist $P > 0$, a diagonal $\Lambda > 0$, $q > 0$, and $K' \in \mathbb{R}^{n\times n_y}$ satisfying (4.11), then the observer (4.7) achieves global exponential convergence of the estimation error to zero for all $t \geq T_0$.*

Proof. For $t \geq T_0$, $y_2 = CWS(V)$ exactly. Consider $\mathcal{V}(e) = e^\top Pe$. Along (4.8),

$$\dot{\mathcal{V}} = -e^\top(PD+DP)\,e + 2\,e^\top P(W-K'CW)\,\delta. \tag{4.12}$$

The sector constraints $\delta_i(\delta_i - \gamma_i\,e_i) \leq 0$ hold for each i. Applying the S-procedure with multipliers $\lambda_i > 0$,

$$\dot{\mathcal{V}} - \sum_{i=1}^{n} 2\lambda_i\,\delta_i(\delta_i - \gamma_i\,e_i) \leq -q\|e\|^2 \tag{4.13}$$

rewrites as

$$\begin{pmatrix} e \\ \delta \end{pmatrix}^\top \begin{pmatrix} -(PD+DP)+qI & P(W-K'CW)+\Gamma\Lambda \\ (W-K'CW)^\top P+\Lambda\Gamma & -2\Lambda \end{pmatrix} \begin{pmatrix} e \\ \delta \end{pmatrix} \leq 0, \tag{4.14}$$

which is guaranteed by (4.11). Thus $\dot{\mathcal{V}} \leq -q\|e\|^2$, giving exponential convergence. □

The LMI (4.11) is bilinear in (P, K') due to the product PK'. With the change of variable $R = PK'$, the off-diagonal block becomes $PW - RCW + \Gamma\Lambda$, which is linear in (P, R, Λ), making it a genuine LMI solvable by semidefinite programming. The gain is recovered as $K' = P^{-1}R$. When C selects a single population j, the effective coupling reads $(I - K'C)W = W - K'W_{j\cdot}$. Choosing $K'_{\cdot 1} = W_{\cdot j}/W_{jj}$ (when $W_{jj} \neq 0$) cancels the j-th column of W, so that the measured population no longer contributes to the error dynamics through the nonlinearity. The design then reduces to ensuring that D absorbs the coupling from the remaining unmeasured populations. More generally, when C measures n_y populations, the $n \times n_y$ free parameters in K' can cancel up to n_y columns of W in the error dynamics.

4.4 COMBINED OBSERVER: LINEAR AND NONLINEAR OUTPUT INJECTION

Since both $y = CV$ and $y_2 = CWS(V)$ are available for $t \geq T_0$, one can combine linear and nonlinear output injection, together with a nonlinearity argument modification via $E \in \mathbb{R}^{n\times n_y}$. This yields the combined observer

$$\dot{\hat{V}} = -D\hat{V} + K(y - C\hat{V}) + WS\big(\hat{V} + E(y - C\hat{V})\big) - K'(CWS(\hat{V} + E(y - C\hat{V})) - y_2) + Bu, \tag{4.15}$$

where $y_2 = CWS(V + E(y - CV)) = CWS(V)$ (since $y - CV = 0$), $K \in \mathbb{R}^{n\times n_y}$ is the linear injection gain, $K' \in \mathbb{R}^{n\times n_y}$ is the nonlinear injection gain, and $E \in \mathbb{R}^{n\times n_y}$ modifies the nonlinearity argument.

Defining $\hat{\zeta} = \hat{V} + E(y - C\hat{V})$ and $\zeta = V$, the estimation error $e = \hat{V} - V$ satisfies

$$\dot{e} = -(D + KC)\, e + (W - K'CW)\, \delta, \qquad \hat{\zeta} - \zeta = (I - EC)\, e, \tag{4.16}$$

where $\delta = S(\hat{\zeta}) - S(\zeta)$. The sector constraint $\delta_i(\delta_i - \gamma_i[(I - EC)e]_i) \le 0$ applies component-wise.

Applying the S-procedure 3.15 with heterogeneous slopes $\Gamma = \mathrm{diag}(\gamma_1, \ldots, \gamma_n)$ and diagonal multiplier $\Lambda > 0$, the combined LMI reads

$$\boxed{\begin{pmatrix} \mathrm{He}\{P(-D - KC)\} + qI & P(W - K'CW) + \Gamma\,(I - EC)^\top \Lambda \\ (W - K'CW)^\top P + \Lambda\,\Gamma\,(I - EC) & -2\Lambda \end{pmatrix} \preceq 0. \tag{4.17}}$$

Theorem 4.4.1. *If there exist $P > 0$, a diagonal $\Lambda > 0$, $q > 0$, and matrices $K, K' \in \mathbb{R}^{n\times n_y}$, $E \in \mathbb{R}^{n\times n_y}$ satisfying (4.17), then the observer (4.15) achieves global exponential convergence of the estimation error to zero for all $t \ge T_0$.*

Proof. For $t \ge T_0$, the sliding-mode reconstruction of Proposition 4.1 guarantees that $y_2(t) = CWS(V(t))$ exactly. Since $y - CV = 0$, the argument substitution gives $\hat{\zeta} - \zeta = (I - EC)e$ and $y_2 = CWS(\zeta)$, so the error dynamics (4.16) hold exactly.

Consider the Lyapunov function $\mathcal{V}(e) = e^\top P e$ with $P > 0$. Its time derivative along the error dynamics (4.16) is

$$\dot{\mathcal{V}} = e^\top \mathrm{He}\{P(-D - KC)\}\, e + 2\, e^\top P(W - K'CW)\, \delta. \tag{4.18}$$

Since S is component-wise slope-restricted in $[0, \gamma_i]$ and $\hat{\zeta}_i - \zeta_i = [(I - EC)e]_i$, the sector constraint yields

$$\delta_i\big(\delta_i - \gamma_i\,[(I - EC)e]_i\big) \le 0, \quad i = 1, \ldots, n. \tag{4.19}$$

Applying the S-procedure (Lemma 3.15) with multipliers $\lambda_i > 0$ collected in $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, a sufficient condition for $\dot{\mathcal{V}} \le -q\|e\|^2$ is

$$\dot{\mathcal{V}} - \sum_{i=1}^{n} 2\lambda_i\, \delta_i\big(\delta_i - \gamma_i\,[(I - EC)e]_i\big) \le -q\|e\|^2 \quad \forall\,(e, \delta). \tag{4.20}$$

Expanding the left-hand side: the term $-2\lambda_i\delta_i^2$ produces the -2Λ block in the $(2, 2)$ position, while the cross-terms $2\lambda_i\gamma_i\delta_i[(I - EC)e]_i$ produce the off-diagonal blocks $\Gamma(I - EC)^\top\Lambda$. The full expression rewrites as

$$\begin{pmatrix} e \\ \delta \end{pmatrix}^\top \begin{pmatrix} \mathrm{He}\{P(-D - KC)\} + qI & P(W - K'CW) + \Gamma\,(I - EC)^\top \Lambda \\ (W - K'CW)^\top P + \Lambda\,\Gamma\,(I - EC) & -2\Lambda \end{pmatrix} \begin{pmatrix} e \\ \delta \end{pmatrix} \le 0, \tag{4.21}$$

which is guaranteed by (4.17). □

Remark 4.4 (Global sector nonlinearities)**.** *When S is not component-wise but satisfies the global sector condition $0 \preceq \partial_\zeta S \preceq \gamma I$ (Proposition 3.17), the n independent constraints collapse into one. Setting $\Gamma = \gamma I$ and $\Lambda = \lambda I$ (scalar $\lambda > 0$) in (4.17) gives a valid LMI with a single S-procedure multiplier. The Galerkin nonlinearity S of Section 3.2 falls into this case.*

The LMI (4.17) specializes to the previous cases:

1. $K' = 0$, $E = 0$: recovers the standard LMI (3.52) with linear injection only.
2. $K = 0$, $E = 0$: recovers the nonlinear injection LMI (4.11) with the $(1,1)$ block $-(PD + DP) + qI$ decoupled from the gain.
3. $K = 0$, $K' = 0$: the E matrix alone modifies the nonlinearity argument, as in [75].

The three gains (K, K', E) act on complementary aspects of the error dynamics: K damps the linear part, K' reshapes the nonlinearity coupling, and E modifies the argument at which the sector constraint is evaluated. With the changes of variable $R_1 = PK$ and $R_2 = PK'$, the terms PK and PK' are linearized. When $E = 0$, the full LMI (4.17) is linear in (P, R_1, R_2, Λ) and can be solved directly by semidefinite programming. When $E \neq 0$, the product EC in $(I - EC)^\top \Lambda$ introduces a bilinearity with Λ; this can be handled by fixing Λ (or E) and alternating.

4.4.1 *The complete observer*

For the sake of completeness, we write the full observer with the super-twisting reconstruction of y_2 and the combined linear and nonlinear injection.

$$\begin{cases} \dot{\hat{z}}_{1,j} = -\lambda\, \hat{z}_{1,j} + \hat{z}_{2,j} + [CBu]_j + L_j \lceil y_j - \hat{z}_{1,j} \rfloor^{1/2}, \\ \dot{\hat{z}}_{2,j} \in L_j^2 \ \mathrm{sign}(y_j - \hat{z}_{1,j}), \qquad j = 1, \dots, n_y, \\ \dot{\hat{V}} = -D\hat{V} + K(y - C\hat{V}) + WS(\hat{\zeta}) - K'(CWS(\hat{\zeta}) - \hat{z}_2) + Bu, \end{cases} \tag{4.22}$$

where $\hat{\zeta} = \hat{V} + E(y - C\hat{V})$. For $t \geq T_0$, the super-twisting observers reconstruct $y_2 = CWS(V)$ exactly for adequate gains $L_j > 0$ (see Proposition 4.1), so the combined observer reduces to the one of Theorem 4.4.1 and enjoys exponential convergence should the LMI be verified.

4.5 COMPARISON WITH THE STANDARD LUR'E OBSERVER ON A CLASSICAL WILSON-COWAN MODEL

The standard application of the general LMI approach to Lur'e systems applied to Wilson-Cowan equations leads to the LMI (3.52), recalled here for convenience:

$$\begin{pmatrix} \mathrm{He}\{P(-D - KC)\} + qI & PW + \gamma\,(I - EC)^\top \Lambda \\ W^\top P + \gamma\,\Lambda(I - EC) & -2\Lambda \end{pmatrix} \preceq 0. \tag{4.23}$$

When both channels are combined in the LMI (4.17), the feasibility region is at least as large as either formulation alone, since setting $K' = 0$ recovers the standard LMI and setting $K = 0$

recovers the pure nonlinear injection LMI. The additional degrees of freedom come at the cost of a finite-time transient T_0 during which the super-twisting observers reconstruct y_2, and of the requirement that solutions remain in a known compact set (to guarantee boundedness of $\dot{z}_2$). In the noise-free setting, the reconstruction is exact and the subsequent convergence is exponential.

In the single-output case $C = e_j^\top$, the gain K' can cancel the j-th column of W in the error dynamics, removing the measured population from the nonlinear coupling. The LMI then only needs D to dominate the coupling from the $n-1$ unmeasured populations, a condition that may be satisfied even when the standard LMI is infeasible.

Beyond feasibility, the two formulations differ in the magnitude of the resulting observer gains as the coupling strength grows. To illustrate this, we compare the classical LMI (solving for K alone, with $K' = 0$) and the combined LMI (4.17) (solving for both K and K' simultaneously, with $E = 0$) on a Wilson–Cowan model [168, 167] of the form

$$\begin{cases} \dot{V}_e = -\lambda V_e + W_{ee} S(V_e) - W_{ei} S(V_i) + u_e, \\ \dot{V}_i = -\lambda V_i + W_{ie} S(V_e) - W_{ii} S(V_i) + u_i, \\ y = CV, \end{cases} \tag{4.24}$$

with $n_b = 3$ excitatory and 3 inhibitory populations ($n = 6$), decay rate $\lambda = 1$, and a base connectivity matrix

$$W = \begin{pmatrix} W_{ee} & W_{ei} \\ W_{ie} & W_{ii} \end{pmatrix}, \quad W_s = s\,W, \tag{4.25}$$

with $W_{ee}, W_{ie} > 0$ and $W_{ei}, W_{ii} < 0$ element-wise and $s > 0$ a scaling factor controlling the overall coupling strength. The measurement matrix $C \in \mathbb{R}^{3\times 6}$ has the form $y_j = \alpha_j (V_e)_j + \beta_j (V_i)_j$ with random positive coefficients. This represent the measurement when an electrode is put on each population measuring an combination of the excitatory and inhibitory activity [68]. The sigmoid S is a standard logistic function with slope $\gamma = 1$. Table 4.1 reports the gain norms returned by the classical and combined LMIs for increasing scaling factor s for the above system with matrix W_s instead of W.

Both LMIs are feasible across all scaling factors with convergence rate $q^* = 50$. The classical gain norm $\|K\|$ grows from 6.6 at $s = 0.5$ to 59.3 at $s = 10$, reflecting the structural role of K in the standard LMI. It must simultaneously provide damping and compensate for the full coupling PW, leading to gains that scale with $\|W\|$. When both channels are available, the combined LMI distributes the effort between K and K'. The linear gain $\|K\|$ stays in the range $[3.5, 5.8]$, while the nonlinear injection gain $\|K'\|$ remains below 2.2 for all values of s. In practice, large gains amplify noise and discretization errors, making the classical observer increasingly fragile as $\|W\|$ grows, even when the LMI is technically feasible.

4.5.1 *Numerical illustration*

We illustrate the observer performance on the same Wilson–Cowan network, scaling the connectivity by $s = 5$ to give $\|W\| = 22.0$. In all simulations we set $E = 0$, so that the combined

s	$\|W_s\|$	$\|K\|_{\text{class.}}$	$\|K\|_{\text{comb.}}$	$\|K'\|_{\text{comb.}}$
0.5	2.2	6.6	3.8	2.12
1.0	4.4	9.5	3.7	1.64
2.0	8.8	15.3	3.5	1.42
3.0	13.2	20.1	3.4	1.36
5.0	22.0	30.3	3.8	1.31
7.0	30.8	40.5	4.7	1.27
10.0	44.1	59.3	5.8	1.24

Table 4.1: Observer gain norms as a function of the coupling strength scaling factor s, for a block Wilson–Cowan model with $n = 6$, $\lambda = 1$, $n_y = 3$. The classical LMI solves for K alone, while the combined LMI solves for both K and K' simultaneously. Both LMIs are feasible across all tested values of s, but the classical gain $\|K\|$ grows linearly with $\|W\|$ while the combined gains remain bounded.

LMI (4.17) is a genuine LMI in the variables (P, R_1, R_2, Λ). The classical LMI ($K' = 0$) returns $\|K\| = 30.3$, while the combined LMI returns $\|K\| = 3.8$ and $\|K'\| = 1.3$.

To evaluate robustness, we perturb the system with Gaussian noise on both the dynamics and the measurements simultaneously:

$$\dot{V} = -DV + WS(V) + u + \sigma_p\, \xi(t), \qquad y = CV + \sigma_m\, \eta(t), \qquad y_2 = CWS(V) + \sigma_m\, \eta_2(t),$$

where $\xi \sim \mathcal{N}(0, I_n)$ is a process noise, $\eta, \eta_2 \sim \mathcal{N}(0, I_{n_y})$ are measurement noises, and both noise sources are present in every simulation. The classical observer corrects through $K(y - C\hat{V})$ alone, while the combined observer uses both $K(y - C\hat{V})$ and $K'(CWS(\hat{V}) - y_2)$.

Table 4.2 reports the root-mean-square estimation error $\|e\|_{\text{rms}}$ at steady state, computed over the interval $[24, 30]$ of a simulation of length $T = 30$ with Euler discretisation at step $\Delta t = 10^{-3}$, for equal noise intensities $\sigma_p = \sigma_m = \sigma$.

σ	Classical $\|e\|$	Combined $\|e\|$	Ratio
10^{-3}	4.18×10^{-3}	1.92×10^{-3}	2.17
5×10^{-3}	2.09×10^{-2}	9.61×10^{-3}	2.17
10^{-2}	4.18×10^{-2}	1.92×10^{-2}	2.17
5×10^{-2}	2.09×10^{-1}	9.61×10^{-2}	2.17
10^{-1}	4.17×10^{-1}	1.92×10^{-1}	2.17

Table 4.2: Estimation error at steady state under simultaneous process and measurement noise of equal intensity $\sigma_p = \sigma_m = \sigma$, with $s = 5$. The combined observer achieves a $2.17\times$ lower error across all intensities.

The error scales linearly with σ for both observers, with a constant ratio of 2.17. This reflects the error dynamics structure: measurement noise enters through the correction terms $K(y - C\hat{V})$ and $K'(CWS(\hat{V}) - y_2)$. Since $\|K\|_{\text{class.}} = 30.3$ while $\|K\|_{\text{comb.}} = 3.8$ and $\|K'\|_{\text{comb.}} = 1.3$, the classical observer amplifies measurement noise by an order of magnitude more.

Figure 4.1 shows the estimation error trajectory $\|e(t)\|$ for both observers under noise intensity $\sigma = 0.01$. Both observers converge within the first few seconds, after which the classical observer fluctuates around a steady-state error roughly twice that of the combined observer.

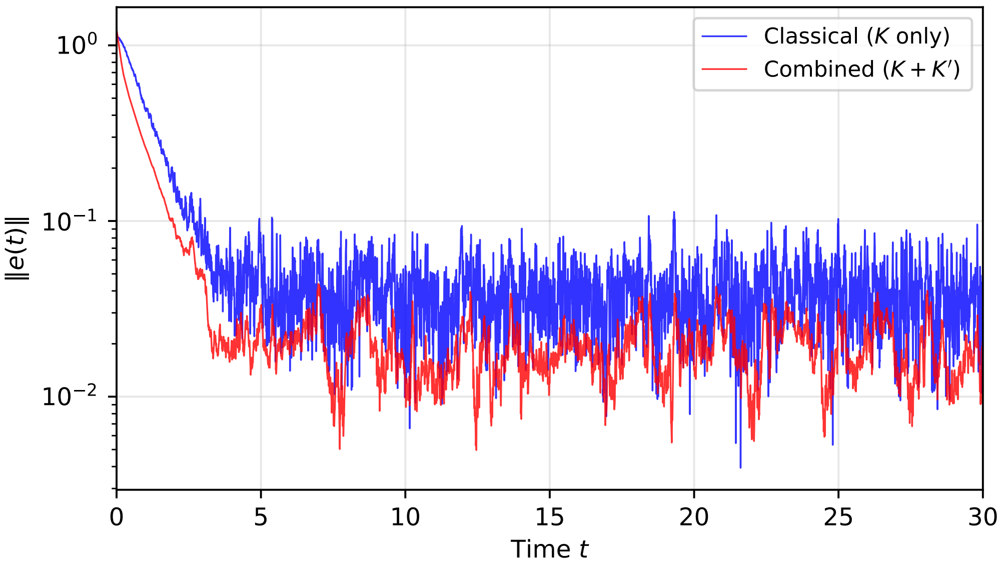


Figure 4.1: Estimation error $\|e(t)\|$ for the classical and combined observers under simultaneous process and measurement noise with $\sigma_p = \sigma_m = 0.01$ and $s = 5$ ($\|W\| = 22.0$). The combined observer achieves a lower steady-state error floor due to its smaller gain norms.

Crucially, the gain norms of the combined observer remain bounded as $\|W\|$ increases (Table 4.1), whereas the classical gain $\|K\|$ grows without bound, eventually making the classical observer impractical regardless of the noise scenario.

4.6 A GENERAL NMM VARIANT VIA AN ENLARGED VIRTUAL OUTPUT

The reconstruction of $y_2 = CWS(V)$ in Section 4.2 requires the uniform-decay assumption $D = \lambda I_n$ (or, equivalently, that $\ker C$ be D-invariant) so that the measurement equation $\dot{y} = -D_y y + CWS(V) + CBu$ is well-posed. We now show that the same approach can be applied to the more general system

$$\begin{cases} \dot{V} = AV + WS(V) + Bu, \\ y = CV, \end{cases} \tag{4.26}$$

with $A \in \mathbb{R}^{n \times n}$ arbitrary, by reconstructing the enlarged virtual output

$$y_2 := CAV + CWS(V). \tag{4.27}$$

Differentiating y along (4.26) gives $\dot{y} = y_2 + CBu$, so y_2 appears as the unknown input of a pure integrator with known forcing CBu. Running n_y parallel super-twisting observers as in (4.5), with the drift $-\lambda \hat{z}_{1,j}$ removed, reconstructs $y_2(t)$ exactly in finite time T_0 as soon as $\dot{z}_2$ is bounded, which holds whenever solutions remain in a compact set.

Consider the observer

$$\dot{\hat{V}} = A\hat{V} + WS(\hat{V}) + Bu + K(y - C\hat{V}) - K'\big(CA\hat{V} + CWS(\hat{V}) - y_2\big), \tag{4.28}$$

with $K, K' \in \mathbb{R}^{n\times n_y}$. The estimation error $e = \hat{V} - V$ satisfies

$$\dot{e} = \big[(I - K'C)A - KC\big]\, e + (I - K'C)W\,\delta, \qquad \delta = S(\hat{V}) - S(V), \tag{4.29}$$

where $\delta_i = \Delta_i e_i$ with $\Delta_i \in [0, \gamma_i]$ by the mean value theorem. The gain K' now acts simultaneously on the linear part and on the nonlinear coupling, while K retains its role of injecting the linear output error.

Remark 4.5. *A further generalization could have been the introduction of a separate gain E modifying the nonlinearity argument, as in (4.15). We choose not to include this degree of freedom here to ease the clarity of this, as was done in the previous section.*

Applying the S-procedure with $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) > 0$ and $\Gamma = \operatorname{diag}(\gamma_1, \ldots, \gamma_n)$, the contraction condition $\dot{\mathcal{V}} \le -q\|e\|^2$ for $\mathcal{V}(e) = e^\top P e$ becomes the matrix inequality

$$\boxed{\begin{pmatrix} \operatorname{He}\{PA - PK'CA - PKC\} + qI & P(W - K'CW) + \Gamma\,\Lambda \\ (W - K'CW)^\top P + \Lambda\,\Gamma & -2\Lambda \end{pmatrix} \le 0.} \tag{4.30}$$

With the changes of variable $R_1 = PK$ and $R_2 = PK'$, the $(1,1)$ block becomes $\operatorname{He}\{PA - R_2CA - R_1C\} + qI$ and the off-diagonal block becomes $PW - R_2CW + \Gamma\Lambda$, both linear in (P, R_1, R_2, Λ). The gains are recovered as $K = P^{-1}R_1$ and $K' = P^{-1}R_2$.

Theorem 4.6.1. *If there exist $P > 0$, a diagonal $\Lambda > 0$, $q > 0$, and matrices $K, K' \in \mathbb{R}^{n\times n_y}$ satisfying (4.30), then the observer (4.28) achieves global exponential convergence of the estimation error to zero for all $t \ge T_0$.*

Proof. For $t \ge T_0$, the parallel super-twisting observers reconstruct $y_2 = CAV + CWS(V)$ exactly, so the error dynamics (4.29) hold without residual term. Consider $\mathcal{V}(e) = e^\top P e$. Along (4.29),

$$\dot{\mathcal{V}} = e^\top \operatorname{He}\{P[(I - K'C)A - KC]\}\, e + 2\, e^\top P(I - K'C)W\,\delta. \tag{4.31}$$

The sector constraints $\delta_i(\delta_i - \gamma_i e_i) \le 0$ hold for each i. Applying the S-procedure (Lemma 3.15) with multipliers $\lambda_i > 0$,

$$\dot{\mathcal{V}} - \sum_{i=1}^{n} 2\lambda_i\, \delta_i(\delta_i - \gamma_i e_i) \le -q\|e\|^2 \tag{4.32}$$

rewrites as

$$\begin{pmatrix} e \\ \delta \end{pmatrix}^\top \begin{pmatrix} \operatorname{He}\{PA - PK'CA - PKC\} + qI & P(W - K'CW) + \Gamma\Lambda \\ (W - K'CW)^\top P + \Lambda\Gamma & -2\Lambda \end{pmatrix} \begin{pmatrix} e \\ \delta \end{pmatrix} \le 0, \tag{4.33}$$

which is guaranteed by (4.30). Thus $\dot{\mathcal{V}} \le -q\|e\|^2$, giving exponential convergence. □

When $A = -D$ with $D = \lambda I_n$, the term $CA = -\lambda C$ and the $(1,1)$ block of (4.30) reduces to $\mathrm{He}\{P(-\lambda I + \lambda K'C - KC)\} + qI$. Setting $\tilde{K} = K - \lambda K'$ recovers exactly the $(1,1)$ block of (4.17) (with $E = 0$), so (4.17) is the special case of (4.30) obtained when the linear part is a scalar multiple of the identity. For a general diagonal D, the term $CA = -CD$ is no longer proportional to C and the gain K' provides a genuinely new degree of freedom in the $(1,1)$ block, which the previous formulation could not exploit.

The price of this generality is the reconstruction step. The signal $z_2 = CAV + CWS(V)$ has a derivative whose magnitude scales with $\|A\|^2$ in addition to $\|W\|$, so the super-twisting threshold gain L_j^* is larger than for $z_2 = CWS(V)$ alone. In a noisy setting this translates into a stronger amplification of measurement noise during the reconstruction phase, which has to be balanced against the larger feasibility region of (4.30).

4.7 NEURAL FIELDS ESTIMATION FROM (LFP) MEASUREMENT CHANNELS

We illustrate the combined observer of Section 4.4 on the estimation of the state of a one-dimensional Amari field from a small number of Local Field Potential (LFP)-like channels, in the spirit of [83, 63].

Recalling the finite dimensional reduction introduced in Section 3.2, we propose a Galerkin/Reduced model that uses $n = 33$ Gaussian basis functions of width $\sigma_\phi = 0.19$, uniformly placed on the periodic interval $\Omega = [0, L]$, $L = 10$, and normalised so that $(\Gamma_{\mathrm{gram}})_{ii} = 1$. The field parameters are taken for theoretical purposes: $\tau = 1$, sigmoidal firing rate $f(x) = (1 + e^{-x})^{-1}$, Mexican-hat kernel $w(r) = a_e\, e^{-r^2/2\sigma_e^2} - a_i\, e^{-r^2/2\sigma_i^2}$ with $(a_e, a_i, \sigma_e, \sigma_i) = (4.0, 1.45, 0.72, 1.6)$. The reduced model is

$$\dot{V} = -\tau^{-1} V + WS(V) + Bu, \qquad y = CV, \tag{4.34}$$

with $W = \tau^{-1}\Gamma_{\mathrm{gram}}^{-1}\widetilde{W}$ and $C_{ji} = \langle m(\cdot - r_j), \phi_i\rangle$. A moving Gaussian stimulus $I(r,t) = 1.5\, e^{-(r - r_c(t))^2/2\sigma_s^2}$, $\sigma_s = 1.0$, $r_c(t) = (0.5\, t) \bmod L$ sweeps the full domain in $T = 20$ time units. Measurements follow (3.7) with Gaussian electrode kernels ($\sigma_m = 0.40$) at $n_y = 16$ evenly spaced sensors.

The nonlinearity S in the reduced model (recall that in this case it doesn't act component-wise) satisfies $0 \preceq \partial_V S \preceq f'_{\max}\, \Gamma_{\mathrm{gram}}$ by Lemma 3.5, so $\|\partial_V S(V)\| \le \gamma$ with

$$\gamma \;=\; f'_{\max}\, \lambda_{\max}(\Gamma_{\mathrm{gram}}) \;=\; 0.562. \tag{4.35}$$

4.7.1 *Observer design*

Since $A = -\tau^{-1} I_n$ with $\lambda = \tau^{-1} = 1$, we apply Theorem 4.4.1 with $E = 0$. By Lemma 3.5, S satisfies $0 \preceq \partial_V S \preceq \gamma I$, so the global sector condition holds and we set $\Gamma = \gamma I$, $\Lambda = \lambda I$ (scalar $\lambda > 0$). Bisecting over $q \in [0, 10]$ returns

$$q^\star \approx 10.0, \qquad \|K\| = 11.27, \qquad \|K'\| = 134.41.$$

With sixteen sensors and thirty-three basis functions, the LMI is feasible up to $q^\star \approx 10$. The large value of $\|K'\|$ reflects the rank deficit of CW: with $n_y = 16$ sensors the matrix CW has rank at most 16, leaving $n - \text{rank}(CW) = 17$ directions of W that K' cannot cancel via output injection, which forces a large compensating gain. This is benign in noise-free simulations but amplifies measurement noise, as quantified in Table 4.3.

4.7.2 *Numerical results*

The reference field $\phi(r)^\top V(t)$ is the n-dimensional Galerkin ODE integrated directly over $[0, T]$ with $\Delta t = 10^{-3}$ and observer initialised at $\hat{V}(0) = \mathbf{1}_n$ (offset $\|V(0) - \hat{V}(0)\| = \sqrt{n}$). Measurements $y(t) = CV(t) + \sigma\xi(t)$ and $y_2(t) = CWS(V(t)) + \sigma\xi'(t)$ ($\xi, \xi' \sim \mathcal{N}(0, I)$ i.i.d.) are consistent with the reduced model up to additive noise $\sigma = 0.1$. In this simulation, $y_2(t) = CWS(V(t))$ is computed from the reduced-model coefficients $V(t)$ and we assume that the reconstruction step has already been achieved. In the noise-free case the error decays exponentially to $\approx 7 \times 10^{-4}$; with $\sigma = 0.1$ the final field error is $\approx$0.077, driven by noise amplification through the gain $\|K'\|$.

Figure 4.2 shows the space-time evolution of the reference traveling bump, which completes one full loop over $T = 20$.

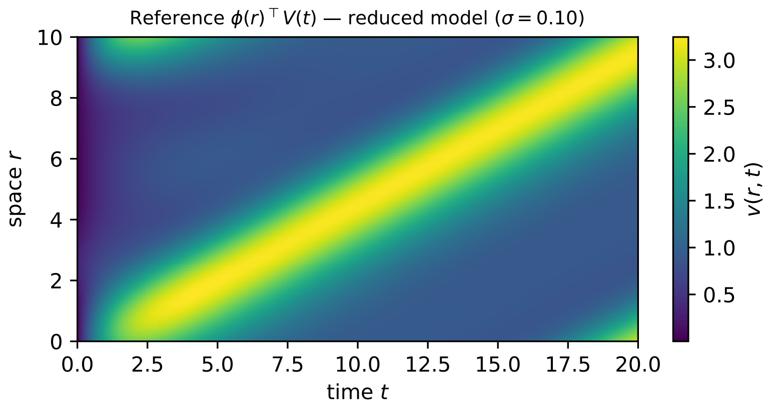


Figure 4.2: Reference field $\phi(r)^\top V(t)$ from the $n = 33$ reduced Galerkin ODE. The traveling bump completes one full loop over $T = 20$.

Figure 4.3 shows the field error $\|v - \hat{v}\|_{L^2}$ over time ($\sigma = 0.1$): the error converges to a noise-driven floor.

Figure 4.4 shows the sixteen measured channels alongside the estimated space-time field $\hat{v}(r, t)$. Each channel panel captures the traveling bump's passage as a local amplitude peak; the estimated field on the right confirms correct spatial reconstruction.

Table 4.3 quantifies the sensitivity to measurement noise.

Figure 4.3: Field reconstruction error $\|v - \hat{v}\|_{L^2}$ ($\sigma = 0.1$, $n = 33$, $n_y = 16$).

Table 4.3: Final field error $\|v - \hat{v}\|_{L^2}$ at $t = T$ for increasing noise level σ. Parameters: $n = 33$, $n_y = 16$, $q^\star \approx 10$.

σ	Reduced model
0.00	7×10^{-4}
0.01	3.9×10^{-2}
0.02	7.7×10^{-2}
0.05	0.194
0.10	0.387

4.8 CONCLUDING REMARKS ON THE OBSERVABILITY OF NEURAL MASS MODELS

Several observer designs have been presented for the neural mass models (3.1). The high-gain, homogeneous, cascade, and interval observer designs all require structural conditions on W or on the dissipation term D that are rarely satisfied in practice. LMI-based Lur'e observers handle the general case: observability reduces to feasibility of an LMI, which can be checked numerically for any connectivity matrix and measurement configuration. A further practical advantage is that the Lyapunov structure yields ISS directly, so bounded disturbances produce a bounded steady-state error with an explicit gain.

In this chapter, we improved on the LMI-based observer introduced in Chapter 3.6 by proposing a preliminary step that reconstructs new outputs of the form $y_2 = CAV + CWS(V)$ via a super-twisting sliding-mode observer. Taking this auxiliary output into account, we proposed the LMI (4.30), which is less conservative than the standard one (3.52): the gain K' acts directly on the coupling term in the error dynamics, so the feasibility condition is weaker and the resulting gains remain bounded as the coupling strength grows. We validated the approach on field reconstruction from partial LFP measurements. The neural field was well approximated by a reduced model of dimension $n = 33$; on this model, the observer (4.22)

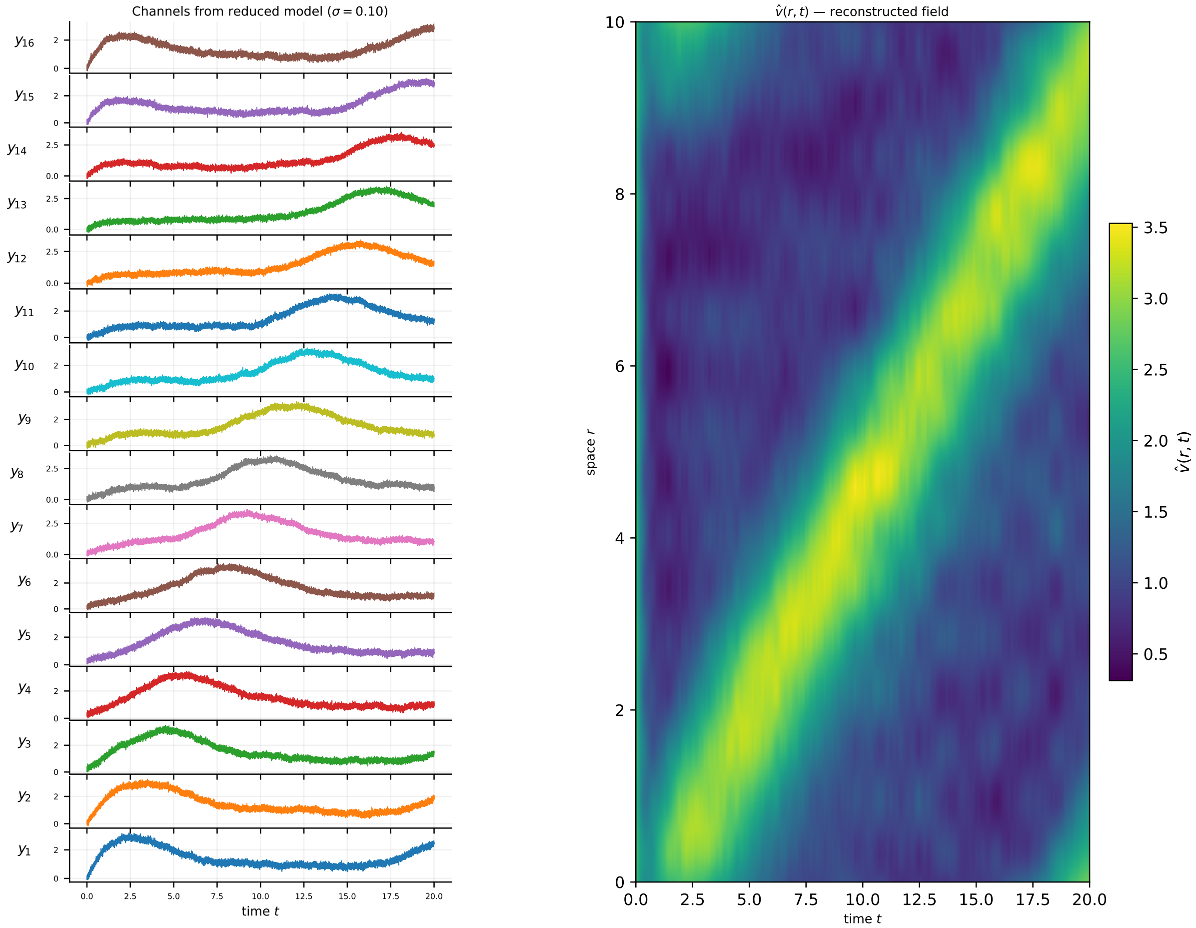


Figure 4.4: Measured channels $y_j(t)$, $j = 1, \ldots, 16$ (left) and estimated field $\hat{v}(r,t)$ (right). The observer reconstructs the traveling bump from 16 sensors ($\sigma = 0.1$).

converges at rate $q^\star \approx 10$ with only $m = 16$ sensors, and the residual error scales predictably with measurement noise. For these reason, the observer proposed in this chapter (4.22) is a promising observer for neural mass models in practice.

5

ACTIVITY ESTIMATION VIA DISTRIBUTED MEASUREMENTS IN AN ORIENTATION SENSITIVE NEURAL FIELDS MODEL OF THE VISUAL CORTEX (PUBLISHED MCSS 2025)

This chapter reproduces the article [8], published in *Mathematical Control and Signals, and Systems*.

This paper investigates the online estimation of neural activity within the primary visual cortex (V1) in the framework of observability theory. We focus on a low-dimensional neural fields modeling hypercolumnar activity to describe activity in V1. We utilize the average cortical activity over V1 as measurement. Our contributions include detailing the model's observability singularities and developing a hybrid high-gain observer that achieves, under specific excitation conditions, practical convergence while maintaining asymptotic convergence in cases of biological relevance. The study emphasizes the intrinsic link between the model's non-linear nature and its observability. We also present numerical experiments highlighting the different properties of the observer. [1]

5.1 INTRODUCTION

Having techniques to estimate online the neural activity of certain cortical areas has many potential uses. This can range from monitoring and treating brain health [149] to applications in neuroscience, such as psychology [13], brain-computer interfacing [16], or the analysis of specific neural phenomena such as visual illusions [96]. Recently, this has opened the door to the design of control theory inspired techniques, feedback-loop control such as in the fields brain-machine interface [150] and deep brain stimulation [99]. For instance, stabilization of signals in the brain via feedback control could be used to alleviate Parkinson's disease symptoms, see simulations in [62]. In most practical cases, neural activity can only be partially measured. It may then be crucial to be able to provide efficient and reliable online estimation methods based on measurements [103, 13, 98].

1 The two articles presented in the following chapters both rely on triangular observer structures, mainly the high-gain and homogeneous observers designed on triangular forms. Readers interested in the theoretical background underlying these constructions are referred to Appendix A, which provides a self-contained account of triangular forms, the associated canonical observability notions, and the main observer design methods as well as the main challenges of using such methods.

In the present study, we are interested in models where the neuronal activity is a distribution over space evolving in time according to well known neural fields equations. These models rely on integro-differential equations obtained by averaging the activity of large groups of neurons in the brain. They were introduced in the seminal works [167, 3] and provide a powerful theoretical framework for studying brain activity. See [36] for comparison of different large scale models of neuronal activity, [153] for a review of neural fields models and [38, 52] for a more in depth analysis of these equations. These models have a rich history of application, particularly in the study of the primary visual cortex V1 [33, 19, 38], and in particular for explaining visual illusions [27, 28, 152, 161]. They have also demonstrated the capability to replicate numerous phenomena observed in experimental data, often acquired through voltage-sensitive dye (VSD) imaging, see [33, 125, 133] for example.

The focus of this paper is on state estimation for a low dimensional model of V1 introduced in [33]. This particular model incorporates many characteristic properties of the *ring model* introduced in [19] and demonstrates a noteworthy ability to accommodate qualitatively for experimental data, on orientation and selectivity, as shown in [133]. To reflect the limitations of standard physical measurement in the cortex, as those obtained via electrodes, we consider distributed measurements and in particular the average voltage within the area of interest.

The considered models can be viewed as control systems, with an input (stimulation), a state space (here of dimension 3), representing the full information needed to predict the future *in the model*, and an output, that will be here a scalar function of the state, representing the partial measurement. In this context, it is customary to study observability (i.e. whether or not the measured outputs contain sufficient information about the full state) and then design an asymptotic observer, i.e. a dynamical system who takes measurements from the system as inputs, and outputs an estimate that asymptotically converges to the "real" state. For recent reviews on observability theory and observer design, see *e.g.* the review article [23], the collection of articles [29] or the textbook [87].

From the standpoint of observability, the considered model introduces distinct challenges as a non-linear system. A critical issue arises from the presence of singular regions where the system may become unobservable (regardless of input). The observability results are contained in Proposition 5.3.1, Proposition 5.3.3, Theorem 5.3.4. First, outside the singular regions, we construct an embedding (of the original 3D state space in $\mathbb{R}^4$) that transforms the system into a canonical form, enabling the design of an observer in this larger space (here adopting a high gain construction, see *e.g.*, [86, 45, 111]) This follows general concepts for the design of observers exposed in [23], but a lot has to be handcrafted. For instance in the present work, we propose a global extension of the dynamics outside the image of the embedding and a uniformly continuous pseudo-inverse of the embedding to recover an estimate of the state. See for example, [24] for a general discussion of this issue. This ensures that the observer operates effectively in regions where observability holds. Second, to handle the singular regions where observability is lost, we introduce a hybridization mechanism that shuts off the part of the observer that becomes singular when crossing these regions to prevent blow-up of the estimates. Theorem 5.3.8, our main theoretical contribution regarding observer convergence, states that, under some excitation condition of the input, the resulting observer guarantees asymptotic convergence if the state trajectory remains in a safe region of the state space and practical convergence in general.

The application to neural fields model of observability theory is quite recent. Up to our knowledge, the only study in this direction is [40] which proposes an adaptive observer for a certain neural fields equation. In that paper, the authors consider two interconnected cortical zones, one of which is thoroughly measured. Their observer design is then based on the contractivity of the dynamics of the un-measured zone and a persistence of excitation condition.

The paper is structured as follows. In Section 5.2 we present the neural fields model of V1 that is the focus of the observer study and establish the precise 3-D model used in the sequel. In Section 5.3 we present the results we obtain in the paper, first on observability properties of the system, then on our tunable observer, whose precise design is to be found in Sections 5.5 and 5.6. The technical part of the paper is divided into three sections. First, Section 5.4 contains technical preliminaries. Then in Section 5.5 we prove the main results regarding observability and give the construction of a pseudo-inverse needed in the observer. Finally in Section 5.6 we prove the main properties of the observer. In the last section, Section 5.7, we discuss a numerical simulation of the observer.

NOTATION. We denote any $X \in \mathbb{R}^3$ as $X = (X_0, X_1, X_2)^\top$. Moreover, we let $X_{1:2} \in \mathbb{R}^2$ be the vector composed of the two last elements, i.e., $X_{1:2} = (X_1, X_2)^\top$. We denote also by $f(v, I(t)) = (f_0(v, I(t)), f_1(v, I(t)), f_2(v, I(t)))$ the right hand side of equation (5.1).

5.2 MAIN MODEL OF STUDY

In this section we present the 3-dimensional model on which the observer will be constructed. We consider the following control system with state $v = (v_0, v_1, v_2) \in \mathbb{R}^3$ and output $y = v_0$

$$\begin{cases} \tau\dot{v}_0 &= -v_0 + J_0 \int_\Omega \sigma(V(r,\theta,v))\frac{P(r)}{\pi}d\theta dr + I_0(t) & := f_0(v, I(t)), \\ \tau\dot{v}_1 &= -v_1 + J_1 \int_\Omega r\cos(2\theta)\sigma(V(r,\theta,v))\frac{P(r)}{\pi}d\theta dr + I_1(t) & := f_1(v, I(t)), \\ \tau\dot{v}_2 &= -v_2 + J_1 \int_\Omega r\sin(2\theta)\sigma(V(r,\theta,v))\frac{P(r)}{\pi}d\theta dr + I_2(t) & := f_2(v, I(t)), \\ y &= v_0. \end{cases} \tag{5.1}$$

Here, τ, J_0 and J_1 are real non-zero parameters of the system. As we need to assume $\tau > 0$ and it does not interfere in our study of the observability, we set $\tau = 1$ from here to ease the proofs. The function $I = (I_0, I_1, I_2)$ is the external input, which we assume to be at least continuous. Regarding the integral term, $\Omega = [0, +\infty) \times [-\pi/2, \pi/2)$ and $P(\cdot)$ is a compactly supported probability density over $[0, +\infty)$. In the integrand, σ is a sigmoidal function to be precised in the following section (Assumption 6.1) and $V(r, \theta, v)$ is given by

$$V(r,\theta,v) = v_0 + rv_1\cos(2\theta) + rv_2\sin(2\theta). \tag{5.2}$$

The output of the system y is the first variable of the state v.

Concerning the existence and the maximal bound of the solutions, we have the following proposition, the proof of which is found in Section 5.4.

Proposition 5.2.1. *Assume $I \in C^0([0,+\infty))$. The dynamics (5.1) admit a unique global solution v defined on $\mathbb{R}$ for every initial condition $v(0) \in \mathbb{R}^3$. Moreover, if $\sup_t \|I\| < \infty$, then there exist $R^* > 0$, such that for any $R \geq R^*$, $B_{\mathbb{R}^3}(0,R)$ is an invariant attracting set by the dynamic f.*

Remark 5.1. *In the case where P is a Dirac mass, (5.1) reduces to the ring model introduced in [19]. Observe also that the assumption of compact support for P is not essential. Indeed, our analysis applies as soon as P admits moments up to the second order.*

Before presenting the primary challenges and key findings related to this system, we outline the derivation of the model in the following section.

5.2.1 *Neural Fields Modelisation*

In this section, following the steps in [33], we present a neural field models of V1 and the necessary steps to obtain the low dimensional model (5.1) on which the observer is constructed.

The goal of the present study is to estimate the post-membrane potential over the visual cortex V1 represented by a bounded open set $\hat{\Omega} \subset \mathbb{R}^2$. We denote by $V(x,t) \in \mathbb{R}$ the average post-membrane potential at a point $x \in \hat{\Omega}$ at a time t. The evolution of V is modeled by a single-layer neural field equation (see *e.g.*, [153])

$$\begin{cases} \tau\partial_t V(x,t) = -V(x,t) + J\cdot\sigma(V(\cdot,t))(x) + I_{\text{ext}}(x,t), \\ V(0) \in L^2(\hat{\Omega}). \end{cases} \tag{5.3}$$

Here, $I_{\text{ext}} \in L^\infty(\mathbb{R}, L^2(\hat{\Omega})) \cap C^0(\mathbb{R}, L^2(\hat{\Omega}))$ denotes the external input, $\tau > 0$ the temporal synaptic constant. The operator J models the neuronal interaction and is given by an integral kernel $j(\cdot,\cdot) \in L^\infty(\hat{\Omega}\times\hat{\Omega})$ through the expression

$$J\cdot u(x) = \int_{\hat{\Omega}} j(x,y)u(y)dy, \qquad u \in L^2(\hat{\Omega}). \tag{5.4}$$

The firing rate function $\sigma : \mathbb{R} \to \mathbb{R}$ models how a population "charges" or "discharges" and is typically chosen to be a smooth sigmoidal function. Equation (6.1) is well defined and has a unique global solution for each initial condition $V(\cdot,0) \in L^2(\hat{\Omega})$ which has the same regularity as I_{ext}. See [38, 162] for instance.

In this work we will assume[2] σ to satisfy the following.

Assumption 5.2.2. *The function σ is a $C^\infty(\mathbb{R},\mathbb{R})$ function that is odd, strictly increasing (i.e., $\sigma' > 0$), convex-concave (i.e., $x\sigma''(x) \leq 0$ for all $x \in \mathbb{R}$), and such that*

$$\lim_{x\to-\infty}\sigma(x) = -1, \quad \sigma(0) = 0, \quad \lim_{x\to+\infty}\sigma(x) = 1, \qquad \max_{\mathbb{R}}\sigma' = \sigma'(0). \tag{5.5}$$

We are interested in the observability and observer design problem for solutions of (6.1). As mentioned, we consider as measurement the cortical activity averaged over the whole domain, as represented by the function

$$h(V(\cdot,t)) = \int_{\hat{\Omega}} V(x,t)\,dx. \tag{5.6}$$

2 See Appendix 5.7, for a discussion on the adjustments required to adapt the ensuing ensuing analysis to the case of strictly positive sigmoids and/or the presence of thresholds h_0, i.e., $\sigma(\cdot - h_0)$.

In the context of our study, the observability problem pertains to determining whether the internal state of a neural field, described by (6.1), can be fully inferred from the cortical activity measurements averaged over the domain, as defined by (6.4). Formally, this problem explores the feasibility of reconstructing the neural field's dynamic state $V(x,t)$ solely from the output $y(t)$. The observer design problem involves creating a system to estimate the internal state of a model based on the output y.

The non-linearity of the sigmoid function σ plays a crucial role in the observability of the system. In the hypothetical case where σ was linear, the dynamics of $t \mapsto h \circ V(\cdot,t)$ will feature only the compositions of $h \circ V(\cdot,t)$ and the system's inherent constants. rendering Equation (6.1) non-observable. It is precisely the non-linear characteristics of σ that enable us to differentiate between potential solutions.

Remark 5.2. *Equation* (6.1) *is what is called in the literature a voltage based model, whereas the original models in [19] and [33] are activity based models. We find this formulation more convenient for mathematical analysis. Moreover, as we can always transition from an activity based model to a voltage-based one via an adequate change of variables, our results can be applied to the former. See Appendix 5.7.*

5.2.2 *Finite-dimensional reduction*

From [33], since the experimental solutions tends to be unimodal, we follow the steps introduced in the article to obtain a reduced model. We are interested in two features encoded by neurons in V1: orientation $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2})$ and selectivity preference $r \in [0,\infty)$ [104]. Mainly, we assume that there exists a polar mapping $\Phi : x \in \hat{\Omega} \mapsto (r_x, \theta_x) \in \Omega := [0,+\infty) \times [-\pi/2, \pi/2)$. See Figure 5.1b. Henceforth, we will work on (6.1) in the coordinates induced by Φ.

The polar mapping Φ induces a change of measure:

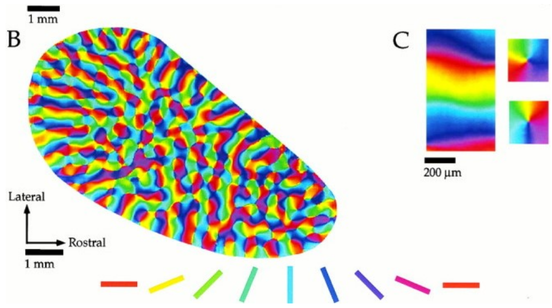
$$dx = P(r,\theta)drd\theta,$$

where $P(r,\theta) = c\,|\det(\partial_{r,\theta}\Phi^{-1})|/|\hat{\Omega}|$ and $c > 0$ is a normalization constant making P a probability density. Following [19] and [33], we suppose that the distribution over orientation preference is uniform, i.e that cortex handle all orientations similarly, obtaining then that $P(r,\theta) = \frac{P(r)}{\pi}$. Since neuronal selectivity ranges in a finite interval, we henceforth assume P to have compact support.

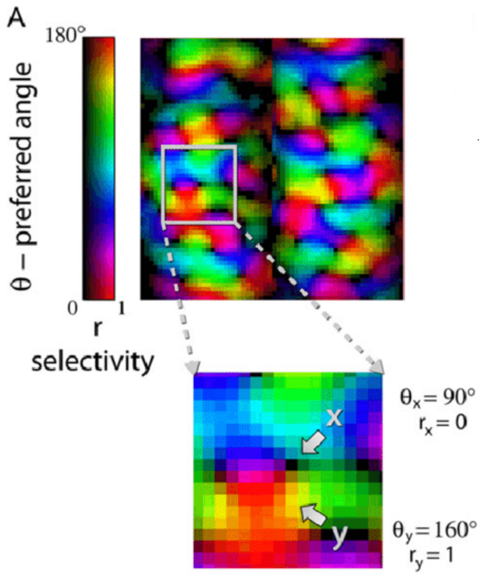
Following the steps in [33] and [162], we look for an operator J in (5.4) such that the experimentally observed cortical stationary states at rest ($I_{\text{ext}} = 0$) are in its range. Since these stationary states tends to be unimodal i.e. a single dominant fourier mode, we can then well approximate the connectivity kernel by :

$$j(x,y) = J_0 + J_1 r_x r_y \cos(2(\theta_x - \theta_y)), \quad J_1 > 0, \quad J_0 \neq 0. \tag{5.7}$$

See [33] and [160][Chapter 10]. Here, the parameters J_0 and J_1 encode, respectively, the strength of global inhibition or excitation and the strength of the connectivity. As J is a compact



(a) Comprehensive orientation preference (OP) map in V1 visualized through optical imaging, depicting the spatial distribution of orientation-selective neurons. See [34].



(b) Orientation preference and selectivity mapping in the visual cortex V1, showing angular preference (top) and selectivity index (bottom) with key orientation angles highlighted. Here the values of the selectivity have been normalized. See [133].

Figure 5.1: Visual cortical mapping of orientation preference and selectivity in V1 region: (a) illustrates a comprehensive view of orientation-selective neuronal distribution through optical imaging, and (b) presents a detailed analysis of orientation preference and selectivity, with emphasis on the normalized selectivity index and key angular orientations.

self-adjoint operator, using that for any $U(\cdot,t) \in L^2(\hat{\Omega})$, $U = U^{\parallel} + U^{\perp}$, where $U^{\parallel} \in \mathrm{range}(J)$, Equation (6.1) splits into

$$\begin{cases} \tau \dot{V}^{\parallel} = -V^{\parallel} + J \cdot \sigma(V^{\parallel} + V^{\perp}) \ + I^{\parallel}, \\ \tau \dot{V}^{\perp} = -V^{\perp} + I^{\perp}, \end{cases} \tag{5.8}$$

with

$$V^{\parallel}(r_x, \theta_x, t) = v_0(t) + v_1(t) r_x \cos(2\theta_x) + v_2(t) r_x \sin(2\theta_x). \tag{5.9}$$

Setting the input such that $I^{\perp}$ is negligible, $V^{\perp}$ converges exponentially towards 0. We henceforth assume $I^{\perp} = 0$ and set $V = V^{\parallel}$.

Finally, observe that the measurement h is recast in the following form

$$h(V(\cdot,t)) = v_0(t) = y(t), \tag{5.10}$$

The variables (v_0, v_1, v_2) follow the dynamical system (5.1) hence the observer design problem introduced in Section 5.2.

Remark 5.3. *Observe that we can also obtain the same measurement by a well placed electrode centered around a pinwheel, i.e. any point $x_0 \in \mathbb{R}^2$ that verifies $r_{x_0} = 0$. Indeed we have using (5.9)*

$$\langle V(\cdot, t), \delta_{x_0} \rangle = v_0(t). \tag{5.11}$$

5.3 STATEMENT OF THE RESULTS

In this section, we present the results achieved in the paper. In a first step we discuss the dynamical properties of the system in connection with observability of the state. Then we propose an estimation method and present its convergence properties under different assumptions.

5.3.1 *Observability*

Here, we investigate the observability properties of (5.1). We first consider the simplest notion of observability, namely to what extent an output $t \mapsto y(t)$, for a prescribed input $t \mapsto I(t)$, can be produced by only one solution $t \mapsto v(t)$. This is called (instantaneous) observability in [23, Section 3.1.2] or [87, Chap. 2, Defn. 1.2].

The following preliminary proposition states that observability *does not* hold when v_0 is zero and gives, away from $\{v_0 = 0\}$, the "maximum information" that can be extracted from the knowledge of the output $y(\cdot)$ without any assumption on the input $I(\cdot)$.

Proposition 5.3.1. *We have the following.*

(i) Let $I \in C^0([0, \infty), \mathbb{R}^3)$ such that $I_0 = 0$. The output $h \circ v = v_0$ of any solution v of (5.1) with initial condition $v(0) = (0, v_1(0), v_2(0))^\top \in \mathbb{R}^3$ is identically null. In this particular case, there exist v and $\hat{v}$ solutions of (5.1) such that

$$h(v(t)) = h(\hat{v}(t)) \text{ and } v(t) \neq \hat{v}(t),\ \forall t \geq 0. \tag{5.12}$$

(ii) Let v and $\tilde{v}$ be two solutions of (5.1) such that $h(v(t)) = h(\tilde{v}(t))$ for all t in an interval $[0, t^*]$, for some $t^* > 0$. Then, for any time $t' \in [0, t^*]$ such that $h(v(t')) \neq 0$ we have

$$|v_{1:2}(t')| = |\tilde{v}_{1:2}(t')| \text{ and } v(t')_{1:2}^\top I_{1:2}(t') = \tilde{v}(t')_{1:2}^\top I_{1:2}(t'). \tag{5.13}$$

Let us now state two more precise results that elaborate on Points *(i)* and *(ii)* above, starting with Point *(i)*, that tells us that observability does not hold in the set $\mathcal{Z}_0 := \{v \in \mathbb{R}^3 : v_0 = 0\}$, and that $\mathcal{Z}_0$ is invariant by the dynamics if the first component of the input is identically zero. Importantly, as the output under consideration is v_0, the value of the output tells us how close the state v is from $\mathcal{Z}_0$. Since we need to build an observer that tolerates trajectories passing through $\mathcal{Z}_0$, Proposition 6.13 below, proved in Section 5.5.1, states that, under the assumption that the first component of the input is bounded from below, trajectories may pass through $\mathcal{Z}_0$ only once and gives an estimate of the time a solution passes "close to" that set. Note that this assumption is often made in the literature, see *e.g.* [162].

Proposition 5.3.2. *Consider an input $I : [0, +\infty) \to \mathbb{R}^3$ such that $I_0(t) \geq c > 0$ for all $t \geq 0$, and a solution $v(\cdot)$ of (5.1) associated to this input. Let*

$$\delta^* = \frac{c}{1 + |J_0|\sigma'(0)}, \tag{5.14}$$

and fix some δ, $0 < \delta < \delta^*$. The first component $v_0(t)$ of the solution vanishes at most one time, and the set of times t such that $|v_0(t)| \leq \delta$ is a time-interval of length no larger than t_δ given by

$$t_\delta = \frac{\delta^*}{c}\frac{2\delta}{\delta^* - \delta}\,. \tag{5.15}$$

Moreover, if there exist a time $t^* \geq 0$ such that $v_0(t^*) \geq \delta$, then $v(t) \geq \delta$ for all times $t \geq t^*$.

Remark 5.4. *If the parameter J_0 is negative and $I_0(t) > c > 0$ the flow of v_0 is always positive when v_0 takes negative values. The proposition above is there solely to treat the case $J_0 > 0$.*

Point *(ii)* in Proposition 5.3.1 states partial observability; it is completed by Proposition 5.3.3 below under proper conditions on $I_{1:2}$.

Proposition 5.3.3. *Consider an input $I : [0, +\infty) \to \mathbb{R}^3$ such that $I_0(t) \geq c > 0$ for all $t \geq 0$. For any $t^* > 0$, the two following assertions are equivalent*

(i) For any two solutions v and $\tilde{v}$ of (5.1) associated to this input,

$$h(v(t)) = h(\tilde{v}(t)) \quad \forall t \in [0, t^*] \quad \Longrightarrow \quad v(t) = \tilde{v}(t) \quad \forall t \in [0, \infty). \tag{5.16}$$

(ii) There exists $t \in [0, t^*]$ such that $I_{1:2}(t) \wedge \dot{I}_{1:2}(t) \neq 0$.

Here, $a \wedge b$ denotes the determinant of the matrix $[a; b]$ for any vector a, b of $\mathbb{R}^2$

The proof of the above results relies on considerations regarding a stronger form of observability on our model, often called *differential obervability* [23, 87], that requires that the value of a certain number of time-derivatives of the output $y = h(v(t))$ at some time t uniquely determines the value of the state $v(t)$ at the same time. The time derivative is given by the differential operator $\mathcal{L}$, also called "total derivative" along the time-varying differential equation (5.1) (see *e.g.*, [23]), that maps a real-valued function $g \in C^k(\mathbb{R}^3 \times \mathbb{R})$ to $\mathcal{L}g \in C^{k-1}(\mathbb{R}^3 \times \mathbb{R}, \mathbb{R})$ defined by

$$\mathcal{L}g(v,t) = f_0(v, I(t))\partial_{v_0} g(v,t) + f_1(v, I(t))\partial_{v_1} g(v,t) + f_2(v, I(t))\partial_{v_2} g(v,t) + \partial_t g(v,t). \tag{5.17}$$

Note that the input $I(\cdot)$ is fixed and sufficiently smooth; a different $I(\cdot)$ defines a different differential operator $\mathcal{L}$. Hinting that the number of time-derivatives needed is no larger than three, we study the time-dependant mapping $v \mapsto T_t(v)$ defined, for all t in $[0, +\infty)$, by

$$T_t(v) = \begin{pmatrix} h(v) \\ \mathcal{L}h\,(v,t) \\ \mathcal{L}^2 h\,(v,t) \\ \mathcal{L}^3 h\,(v,t) \end{pmatrix}, \tag{5.18}$$

The map T_t is often called the observability mapping at time t, indeed its injectivity implies the aforementioned differential observability. The following proposition is our main result on differential observability of (5.1). It is proved in Section 5.5.1. Define the sets

$$\mathcal{Z}_0 = \{v \in \mathbb{R}^3 : v_0 = 0\}, \quad \mathcal{Z}_{1:2} = \{v \in \mathbb{R}^3 \mid v_1 = v_2 = 0\}, \quad \mathcal{Z} := \mathcal{Z}_{1:2} \cup \mathcal{Z}_0 . \tag{5.19}$$

Theorem 5.3.4. *Assume $I \in C^2([0, +\infty), \mathbb{R}^3)$. Let $t \geq 0$ be such that $I_{1:2}(t) \wedge \dot{I}_{1:2}(t) \neq 0$.*

- *The map $T_t : \mathbb{R}^3 \setminus \mathcal{Z}_0 \to \mathbb{R}^4$ is an injection.*
- *The map $T_t : \mathbb{R}^3 \setminus \mathcal{Z} \to \mathbb{R}^4$ is an injective immersion.*

The loss of injectivity in the singular region $\mathcal{Z}_0$ is related to the loss of observability in Proposition 5.3.1. The first point implies that the restriction of T_t to $\mathbb{R}^3 \setminus \mathcal{Z}_0$ has an inverse $T_t(\mathbb{R}^3 \setminus \mathcal{Z}_0) \to \mathbb{R}^3 \setminus \mathcal{Z}_0$; the second point implies that this inverse is differentiable on $T_t(\mathbb{R}^3 \setminus \mathcal{Z})$. In fact this inverse is continuous, but fails to be locally Lipschitz continuous at images of points in $\mathcal{Z}_{1:2}$.

Remark 5.5. *Model (5.1) has been introduced in [33] to accommodate for experimental solutions obtained through VSD. These solutions always verify that*

$$\exists \eta^* > 0, t_{\eta^*} > 0, \quad |v_{1:2}| \geq \eta^*, \ \forall\, t \geq t_{\eta^*}. \tag{5.20}$$

In the next section, we discuss a tunable observer that verifies exponential asymptotic convergence for solutions verifying (5.20), while keeping some form of practical convergence for solutions that do not. We observe that appropriately chosen inputs and parameters may produce solutions that cross the region $\mathcal{Z}_{1:2}$ for arbitrarily large times.

In light of the above discussion, we define the following family of tubular neighborhoods of $\mathcal{Z}$, indexed by two parameters $\delta > 0$, $\eta > 0$:

$$\mathcal{Z}_{\delta,\eta} := \{v \in \mathbb{R}^3 : \ |v_0| < \delta\} \cup \{v \in \mathbb{R}^3 : \ |v_{1:2}| < \eta\}. \tag{5.21}$$

The restriction of the inverse of T_t to a bounded part of $T_t(\mathbb{R}^3 \setminus \mathcal{Z}_{\delta,\eta})$ will have a global Lipschitz constant.

Let us now derive from the above observability conditions an assumption that will be useful for designing an observer in next subsection. We saw in Proposition 6.13 that the trajectories cross the non observavility locus $\{v_0 = 0\}$ at most once if I_0 stays away from zero and in Proposition 5.3.3 that observability occurs when v_0 and $I_{1:2} \wedge \dot{I}_{1:2}$ stay away from zero; in view of these points, we make the following assumption on the input, for the rest of the paper.

Assumption 5.3.5. *The input $t \mapsto I(t)$ is in $C^3([0, +\infty), \mathbb{R}^3)$ and is bounded on $[0, +\infty)$, as well as its time-derivatives of order up to three, and the following lower bounds hold for some positive numbers c and μ:*

$$I_0(t) \geq c > 0 \quad \textit{and} \quad |I_{1:2}(t) \wedge \dot{I}_{1:2}(t)| \geq \mu > 0 \qquad \textit{for all } t \textit{ in } [0, +\infty). \tag{5.22}$$

Remark 5.6. *The same results in this paper can be obtained by assuming $|I_0(t)| \geq c$ rather than $I_0(t) > c$. The positive sign of I_0 in the first relation in (6.7) is chosen to facilitate the exposition.*

Remark 5.7. *The second condition in (6.7) is a stricter version of the standard lower bound for persistence of excitation-type conditions*

$$\exists \mu^*, t^* > 0,\ \forall t_0 > 0,\ \frac{1}{t^*}\int_{t_0}^{t_0+t^*} I_{1:2}(t)^\top I_{1:2}(t)dt \ \succeq\ \mu^* \operatorname{Id}_{2\times 2}, \tag{5.23}$$

where $\operatorname{Id}_{2\times 2}$ *denotes the identity matrix of dimension 2.*

5.3.2 *Observer design*

In this section, we propose an observer based on the high-gain construction [23, 87] that takes into account the observability difficulties encountered in the model. Using the notations of previous section, the mapping $(t, v) \mapsto T_t(v)$ is structured to satisfy the following condition for all $v \in \mathbb{R}^3$ and $t \in \mathbb{R}$:

$$\partial_v T_t(v,t) f(v, I(t)) + \partial_t T(v,t) = AT_t(v,t) \ +\ W\mathcal{L}^4 h(v,t), \tag{5.24}$$

where A and W are matrices defined as:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad W = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{5.25}$$

The next step involves defining a Lipschitz pseudo-inverse $\mathfrak{T}_t$ for T_t at any given time $t \geq 0$. This pseudo-inverse is essential for extracting $v(t)$ from the values $z(t) = T_t(v(t))$ and to make so that only terms depending on z appears in the right-hand side of (5.24). A high-gain can then be designed which estimates the value of $z(t) = T_t(v(t))$ for any solutions v of (5.1) based on its output y. The efficacy of this design hinges on existence of a pseudo-inverse of T_t which is Lipschitz, ensuring the observer's accurate and stable convergence. For this design to be feasible,some form of differential observability is required as mentioned in the previous section. Further details on this construction can be found in [29] for example.

Within the established framework, let us now construct, for each time $t \geq 0$, a globally Lipschitz continuous mapping denoted as $\mathfrak{T}_t$, serving as a left inverse to the observability mapping T_t within the set $\mathcal{Z}^c_{\delta,\eta} \cap B_{\mathbb{R}^3}(0, R)$ [3], where R is a sufficiently large constant, $0 < \eta < R$, and $0 < \delta < \delta^*$, where δ^* is defined in Proposition 6.13. The validity of Proposition 6.14 is invoked to affirm the existence of an arbitrarily large $R > 0$, ensuring that $B_{\mathbb{R}^3}(0, R)$ constitutes an invariant attracting set for the system described by equation (5.1). Subsequently, the use of a projection map alongside the inherent symmetry of the non-linear term in (5.1) is employed to broaden the definition of $\mathfrak{T}_t$ across the entire domain $\mathbb{R}^4$ at any time $t \geq 0$. This is encapsulated within the following theorem, the proof of which is given in Section 5.5.2. See (5.73) for the explicit expression of $\mathfrak{T}_t$.

3 $\mathcal{Z}^c_{\delta,\eta}$ denotes the complement of the set $\mathcal{Z}_{\delta,\eta}$.

Theorem 5.3.6. *Suppose that Assumption 5.3.5 holds. Let $\delta > 0$, $0 < \eta < R$. Then, for any output $y = h \circ v$, the map*

$$\mathfrak{T} : \mathbb{R}^4 \times [0, +\infty) \to \{(u_0, u_1, u_2) \in \mathbb{R}^3, \delta \leq |u_0| \leq R \text{ and } \eta \leq |u_{1:2}| \leq R^2\}, \tag{5.26}$$

explicitely defined in (5.73), such that $\mathfrak{T}_t = \mathfrak{T}(\cdot, t)$ is globally Lipschitz continuous uniformly with respect to t, and satisfies

$$|\mathfrak{T}_t(T_t(u)) - u| \leq \ \eta\chi(u), \quad \forall u \in \mathcal{Z}^c_{\delta,\eta} \cap B_{\mathbb{R}^3}(0, R) \text{ such that } u_0 y(t) > 0,\ |u_0| \geq \delta \tag{5.27}$$

where

$$\chi(v) := \begin{cases} 0 & |v_{1:2}| \geq \eta, \\ 1 & \text{otherwise.} \end{cases} \tag{5.28}$$

Remark 5.8 (on the explicit definition of the map $\mathfrak{T}_t$)**.** *Note that $\mathfrak{T}_t$ is defined in (5.73) as the composition of four other maps; three of them are very explicit while the fourth one id the inverse of an explicit globally defined map (S_t), and computing this inverse can be reduced to inverting a scalar monotonic map, as explained in Remark 5.10.*

We propose now the following observer which takes into account the consideration of Remark 5.5.

Let be $0 < \delta < \delta^*$ where δ^* is defined in Proposition 6.13, $0 < \eta < R$, $l \geq 1$. Let $C = (1, 0, 0, 0) \in \mathbb{R}^{1\times 4}$ and $P_l(i, j) := \frac{(-1)^{i+j}}{l^{i+j-1}} \cdot \frac{(i+j-2)!}{(i-1)!(j-1)!}$. Considering $y = h \circ v$, where v is a solution of (5.1), we study a tunable state observer, piece-wise defined as follows:

$$\begin{cases} \hat{v}(t) = \mathfrak{T}_t(\hat{z}(t)), \quad \dot{\hat{z}}(t) = A\hat{z}(t) + \mathcal{L}^4 h\left(\mathfrak{T}_t(\hat{z}(t)), t\right) W - P_l^{-1} C^\top (C\hat{z}(t) - y(t)), & \text{if } |y(t)| > \delta, \\ \dot{\hat{v}}(t) = f(\hat{v}, t), \quad \hat{z}(t) = T_t(\hat{v}(t)) & \text{if } |y(t)| < \delta. \end{cases} \tag{5.29}$$

This observer is a hybrid system where simple rules, governed by the sign of $|y(t)| - \delta$, switch the dynamics between two time-varying continuous-time dynamical systems (explicit dependence on time comes from I and y, whose evolution does not depend on the observer), with states $\hat{z} \in \mathbb{R}^4$ and $\hat{v} \in \mathbb{R}^3$. Without resorting to a general framework for hybrid systems (see *e.g.* [159]), let us make clear what a solution of this system is.

The initial condition is given by some $\hat{z}(0) \in \mathbb{R}^4$ if $|y(0)| > \delta$ (or $|y(0)| = \delta$ and $|y(t))| > \delta$ for small positive t), and by some $\hat{v}(0) \in \mathbb{R}^3$ if $|y(0)| < \delta$ or $|y(0)| = \delta$ and $|y(t))| < \delta$ for small positive t. We show in Proposition 6.10 that the solutions of the corresponding ODE are well-defined until the first time where $|y(t)| - \delta$ changes sign.

Proposition 6.13 tells us that $|y(t)| - \delta$ changes sign at most twice on $[0, +\infty)$; we have to describe the jumps between state spaces at such a time $t_\star$ (either t_1 or t_2 from Proposition 6.13) to be complete. Let $t_\star$ be the first such time. If $|y(t)| \geq \delta$ on $[0, t_\star]$, $|y(t)| > \delta$ on some $[t_\star - \varepsilon, t_\star)$, the solution $\hat{z}(t)$ of the first line of (6.16) is well defined on some $[0, t_\star + \varepsilon]$ so that it has a limit at $t = t_\star$, that we may denote by $\hat{z}(t_\star^-)$; the solution to the hybrid system (6.16) is continued by initializing the second differential equation with state $\hat{v}$ according to $\hat{v}(t_\star^-) = \mathfrak{T}_t(\hat{z}(t_\star^-))$ on the next interval where $|y(t)| - \delta$ is negative; similarly, if $|y(t)| < \delta$ on some $[t_\star - \varepsilon, t_\star)$ (and

$|y(t)| > \delta$ on some $(t_\star, t_\star + \varepsilon]$), the initialization in the new state space is made according to $\hat{z}(t_\star^-) = T_t(\hat{v}(t_\star^-))$.

Henceforth, with a slight abuse of language, we refer to $\hat{v}$ as the solution of (6.16). We can now state the following, which is proven in Section 5.6

Proposition 5.3.7. *For any $\delta < \delta^*$, with δ^* given by (6.36) in Proposition 6.13, any $0 < \eta < R$, solutions of (6.16) are defined globally, piece-wise continuous, with $\hat{v}$ having at most a single jump.*

Theorem 5.3.8. *Suppose Assumption 5.3.5 holds. For any $R > R^*$, where R^* is defined in Proposition 6.14, for any v solution of (5.1) with output $y = v_0$, and such that $v(0) \in B_{\mathbb{R}^3}(0, R)$, we have the following.*

(i) For any $\delta < \delta^*$, where δ^* is defined by (6.36) in Proposition 6.13, $0 < \eta < R$, there exist l^*, $M > 0$, $L > 0$ such that for any $l \geq 1$, any interval $[t_0, t_f]$ where $|y(t)| \geq \delta$ for $t \in [t_0, t_f]$, we have for any solution $\hat{v}$ of (6.16) with initial conditions $\hat{v}(0) \in \mathbb{R}^3$ and $\hat{z}(0) = T_0(\hat{v}(0))$, it holds for any time $t_0 \leq t < t_f$

$$|v(t) - \hat{v}(t)| \leq \eta\chi(v(t)) + Me^{-(l-l^*)(t-t_0)}\left(l^3|\hat{z}(t_0) - T_{t_0}(v(t_0))| + \int_{t_0}^{t} e^{(l-l^*)(s-t_0)} L\eta\chi(v(s))\, ds\right) \tag{5.30}$$

where χ is defined by (5.28) in Theorem 5.3.6. Moreover, there exists $L_2 > 0$ such that if the the switching interval $\mathcal{S} = \{t \geq 0 : |y(t)| < \delta\}$ is non-empty, letting $t_1 = \inf \mathcal{S}$, it holds

$$|v(t) - \hat{v}(t)| \leq e^{L_2(t-t_1)}|\hat{v}(t_1) - v(t_1)|, \quad t \in \mathcal{S}, \tag{5.31}$$

(ii) Suppose that there exist $\eta^* > 0$ such that $|v_{1:2}(t)| \geq \eta^*$, $\forall t \geq 0$. Then for $\eta = \eta^*$, $0 < \delta < \delta^*$, there exist $l^* > 0$, $M > 0$ and $k > 0$ such that for any $l \geq l^*$ for any solution $\hat{v}$ of (6.16) with initial conditions $\hat{v}(0) \in \mathbb{R}^3$ and $\hat{z}(0) = T_0(\hat{v}(0))$, it holds

$$|\hat{v} - v| \leq k|v(0) - \hat{v}(0)|e^{-(l-l^*)t}, \quad t \geq 0. \tag{5.32}$$

(iii) For any time $t^* > 0$, any compact $\mathcal{K} \subset \mathbb{R}^3$, and any $\varepsilon > 0$, there exist $\delta > 0$, $\eta > 0$, $l > l^*$ sufficiently large such that for any solution $\hat{v}$ of (6.16) with initial conditions $\hat{v}(0) \in \mathcal{K}$ and $\hat{z}(0) = T_0(\hat{v}(0))$, letting $t_2 := \inf\{t \geq 0,\ y(t) = \delta\} = \sup \mathcal{S}$, it holds

$$|\hat{v}(t) - v(t)| \leq \varepsilon, \quad t \in [t^*, t_2] \cup (t_2 + t^*, \infty). \tag{5.33}$$

Let us comment on this theorem. We described in Section 5.3.1 a certain number of observability losses in our model (5.1); these occur for certain behaviors of the input and in some regions of the state space. Our approach has been to make assumptions on the input —this is Assumption 5.3.5— but to then proceed with crafting an ad-hoc versatile observer —this is (6.16), based on the pseudo-inverse whose properties are described by Theorem 5.3.6— capable of achieving tunable exponential convergence when the state trajectory stays away from the identified singularities, or do not approach them "too often", while also ensuring acceptable performance for any other space trajectories. Let us detail and comment these properties, stated formally above.

- Solutions of the observer are defined for all time for any initial condition, for any state trajectory of the system, this is Proposition 6.10.
- It provides tunable exponential convergence when the state trajectory stays in some region $\{|v_{1:2}| > \eta^* > 0\}$; this is point (ii) of the theorem. Note that the hybrid construction allows us to preserve exponential convergence through one possible passage through the non-obervability locus $\{v_0 = 0\}$. We explain in Remark 5.5 to what extent these state trajectories are the ones of interest.One might chose to see this point as the main result. We refer the reader to Section 5.7 for an example of such a trajectory.
- It provides rougher practical convergence when the state trajectory keeps passing in the "bad regions" as time grows; this is point (iii) that gives an estimate that holds *except* on a possible (arbitrarily small) time interval where $|v_0|$ is small.

Point (i) gives an explicit bound on the error taking into account all contributions, without making any assumption on the state trajectory; this bound is possibly difficult to interpret but is necessary for completeness. The other points may be deduced from the general estimations (5.30) and (5.31).

5.4 TECHNICAL PRELIMINARIES

In this section, we prove the main properties of system (5.1) and the observability mapping T introduced in (5.18). We first start by proving the global existence of the solutions of (5.1) and the existence of attracting sets for this system.

Proof of Proposition 6.14. Recall that the function σ is Lipschitz continuous with Lipschitz constant $\sigma'(0)$. Hence, letting $w = (1, r\cos(2\theta), r\sin(2\theta))^\top$, we have for $v, \tilde{v} \in \mathbb{R}^3$:

$$|\sigma(w^\top v) - \sigma(w^\top \tilde{v})| \leq \sigma'(0)\sqrt{1+r^2}\;|v - \tilde{v}|. \tag{5.34}$$

We therefore obtain,

$$|f(v,t) - f(\tilde{v},t)| \leq \left| -1 + \sqrt{(J_0^2 + 2J_1^2)} \int_{[0,\infty)} \sigma'(0)(1+r^2)P(r)dr \right| |v - \hat{v}|, \tag{5.35}$$

which is then bounded uniformly on time as P is compactly supported by assumption. Therefore solutions of (5.1) are well defined for every time.

Let now v be solution of (5.1). Using the fact that σ is bounded by 1, we have that:

$$\dot{v}^\top v = \frac{1}{2}\frac{d|v|^2}{dt} \leq -|v|^2 + \left(\sqrt{(J_0^2 + 2J_1^2)} \int_R (\sqrt{1+r^2}P(r)dr + \sup_t \|I\|\right) |v|,$$

as the integral term is uniformly bounded. Therefore, if $\sup_t \|I\| < \infty$ there exist $R^* > 0$ large enough such that for any $R \geq R^*$, if $|v| > R$ we have $\dot{v}^\top v < 0$. $\square$

To study the observability mapping T, we utilise the $O(2)$ symmetry of the non-linear part of (5.1) and therefore study (5.1) from polar coordinate point of view. This allows us to naturally extend the map T_t into a diffeomorphism at each time t and more easily construct a left inverse.

We now put system (5.1) into polar coordinates. To this effect, we rewrite the system as

$$\dot{v} = -v + \Psi(v) + I, \tag{5.36}$$

where Ψ the non-linear integral part of f. For any $\phi \in [0, 2\pi)$ we have

$$\Psi\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} v\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \Psi(v), \tag{5.37}$$

which is obtained simply through the change of variable $\theta' = \theta + \phi$ in the periodic integrals. To take advantage of the above observation, we introduce the following quantities, for $j, p \in \mathbb{N}$,

$$\Gamma_p^j(v_0, \rho) = \int_\Omega r^j \cos(2\theta)^j \sigma^{(p)}(v_0 + r\rho\cos(2\theta)) P(r) \frac{d\theta dr}{\pi}, \qquad v_0 \in \mathbb{R}, \, \rho \geq 0. \tag{5.38}$$

Here, $\sigma^{(p)}$ denotes the p-th derivative of the real-valued function σ, in particular $\sigma^{(0)} = \sigma$. Letting $v_{1:2} = \rho e_\phi$ with $e_\phi = (\cos\phi, \sin\phi)^\top$, via (5.37) above we obtain

$$\Psi(v) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \Psi\begin{pmatrix} v_0 \\ |v_{1:2}| \\ 0 \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \begin{pmatrix} J_0\Gamma_0^0(v_0, \rho) \\ J_1\Gamma_0^1(v_0, \rho) \\ 0 \end{pmatrix}. \tag{5.39}$$

Here, we used the fact that

$$\int_\Omega r \sin(2\theta)\sigma(v_0 + r\rho\cos(2\theta)) \frac{d\theta}{\pi} P(r) dr = 0, \qquad \forall \rho \geq 0. \tag{5.40}$$

As a consequence, the dynamics (5.1) reads

$$\begin{cases} \dot{v}_0 = -v_0 + J_0\Gamma_0^0(v_0, |v_{1:2}|) + I_0, \\ \dot{v}_{1:2} = -v_{1:2} + J_1\Gamma_0^1(v_0, |v_{1:2}|)\dfrac{v_{1:2}}{|v_{1:2}|} + I_{1:2}. \end{cases} \tag{5.41}$$

The following lemma presents some basic facts about the non-linearities Γ_p^j introduced in (5.38).

Lemma 5.4.1. *The following assertions hold.*

(i) The quantities Γ_p^j are well defined for all $j, p \in \mathbb{N}$ and belong to the class $C^\infty(\mathbb{R} \times [0, \infty), \mathbb{R})$.

(ii) We have the following relations

$$\partial_{v_0}\Gamma_p^j = \Gamma_{p+1}^j \quad \textit{and} \quad \partial_\rho \Gamma_p^j = \Gamma_{p+1}^{j+1}, \qquad \forall j, p \in \mathbb{N}. \tag{5.42}$$

(iii) For any $v_0 \neq 0$, the function $\Gamma_0^0(v_0, \cdot) \in C^\infty([0,\infty), \mathbb{R})$ is injective.

(iv) For any $v_0 \neq 0$ and $\rho > 0$, we have $\Gamma_1^1(v_0, \rho) \neq 0$ and of opposite sign w.r.t. v_0. Moreover, $\Gamma_1^1(v_0, 0) = 0$ for any $v_0 \in \mathbb{R}$.

(v) For any $\rho \geq 0$, we have $\Gamma_0^0(0, \rho) = 0$.

Proof. We start with points *(i)* and *(ii)*. For for $j, p \in \mathbb{N}$, we set

$$\gamma_p^j(v_0, \rho, r, \theta) := \sigma^{(p)}(v_0 + r\rho\cos(2\theta))r^j\cos(2\theta)^j P(r), \quad v_0 \in \mathbb{R}\ \rho \in [0,\infty),\ r \in [0,\infty),\ \theta \in \mathbb{S}^1, \tag{5.43}$$

From our assumptions on σ (see (5.5)), for any fixed θ and almost every r, the function $\gamma_p^j(\cdot, r, \theta)$ is $C^\infty(\mathbb{R} \times [0,\infty), \mathbb{R})$ and verifies

$$\partial_{v_0}\gamma_p^j(\cdot, r, \theta) = \gamma_{p+1}^j(\cdot, r, \theta), \quad \text{and} \quad \partial_\rho\gamma_p^j(\cdot, r, \theta) = \gamma_{p+1}^{j+1}(\cdot, r, \theta). \tag{5.44}$$

Furthermore, using the boundedness of $\sigma^{(p)}$ we see that

$$|\sigma^p(v_0 + r\rho\cos(2\theta))r^j\cos(2\theta)^j P(r)| \leq \|\sigma^{(p)}\|_\infty r^j P(r). \tag{5.45}$$

Using the fact that P is a compactly supported probability density, the above allows to apply Lesbegue dominated convergence theorem, which yields the claim.

Point *(iii)* is a direct consequence of point *(iv)* as *(ii)* implies $\partial_\rho\Gamma_0^0(v_0, \cdot) = \Gamma_1^1(v_0, \cdot)$, so we turn to a proof of *(iv)*.

From Assumption 6.1 on σ, by remarking that σ' is even, strictly increasing on $(-\infty, 0)$ and strictly decreasing on $(0,\infty)$ with $\sigma'(0) = \sup_{\mathbb{R}} \sigma'$, we get that:

$$\begin{cases} \sigma'(v_0 + r\rho U) - \sigma'(v_0 - r\rho U) < 0, & \forall\rho > 0, \quad \text{if } v_0 > 0, \\ \sigma'(v_0 + r\rho U) - \sigma'(v_0 - r\rho U) > 0, & \forall\rho > 0, \quad \text{if } v_0 < 0, \\ \sigma'(v_0 + r\rho U) - \sigma'(v_0 - r\rho U) = 0, & \qquad\qquad \text{if } \rho v_0 = 0. \end{cases} \tag{5.46}$$

We have then, using the fact that r and $P(r)$ are positive, that

$$\begin{cases} \Gamma_1^1(v_0, \rho) < 0, & \forall\rho > 0, \quad \text{if } v_0 > 0, \\ \Gamma_1^1(v_0, \rho) > 0, & \forall\rho > 0, \quad \text{if } v_0 < 0 \\ \Gamma_1^1(v_0, \rho) = 0, & \qquad\qquad \text{if } \rho v_0 = 0. \end{cases} \tag{5.47}$$

This completes the proof of *(iv)*.

Finally, point *(v)* follows at once from direct computations and the parity of σ, which yields

$$\Gamma_0^0(0, \rho) = \int_\Omega \sigma(r\rho\cos(2\theta))P(r)\frac{d\theta dr}{\pi} = 0. \tag{5.48}$$

$\square$

Remark 5.9. *As already mentioned, our results are true without the compact support assumption on P. The results can be obtained with the assumptions that the maximal moment of P, which we denotes m, is greater than 3. Indeed, the above result still holds, supposing that the distribution P has a moment m greater or equal then 2. Indeed by replacing point (i) by the fact that Γ^j_p is well defined for all $j \leq m$ and has regularity $C^{(m-j)}(\mathbb{R} \times [0, \infty), \mathbb{R})$.*

We are now in a position to prove Propositions 5.3.1 and 6.13.

Proof of Proposition 5.3.1. Let v and $\tilde{v}$ be two solutions of (5.1). Let us prove *(i)*. From Lemma 6.12, $\Gamma^0_0 \equiv 0$. Using Equation (5.57), the set $\mathcal{Z}_0 = \{v \in \mathbb{R}^3,\ v_0 = 0\}$ is stable by the dynamic f as $I_0(t) = 0$ for all times $t \geq 0$. Therefore for solution initialised in the set $\mathcal{Z}_0$ the dynamic can be reduced to the close form

$$\dot{v}_{1:2} = f_{1:2}((v_0, v_1, v_2), I(t)), \qquad v_0 = 0, \tag{5.49}$$

which admits a unique solutions for any initial condition $v_{1:2}(0) \in \mathbb{R}^2$ as f is globally Lipschitz continuous from Proposition 6.14. There exist then solutions of (5.1) such that

$$v_0(t) =_0 (t),\ \forall t \geq 0, \text{ and } v(t) \neq \tilde{v}(t),\ \forall\ t \geq 0. \tag{5.50}$$

Let us now prove point *ii*. By assumption, it holds

$$v_0(t) = \tilde{v}_0(t) \quad \text{for} \quad t \in [0, t^*]. \tag{5.51}$$

In particular, $\dot{v}_0 \equiv \dot{\tilde{v}}_0$ on $(0, t^*)$. Hence, using the first equation of (5.41) and the fact that $J_0 \neq 0$, we obtain that

$$\Gamma^0_0(v_0(t), \rho(t)) = \Gamma^0_0(v_0(t), \tilde{\rho}(t)) \quad \text{for} \quad t \in [0, t^*]. \tag{5.52}$$

For $t \in \mathcal{I} = \{t \in [0, t^*] :\ v_0(t) \neq 0\}$, point *(iii)* of Lemma 6.12 guarantees that $\Gamma^0_0(v_0(t), \cdot)$ is injective, and thus that

$$\rho(t) = \tilde{\rho}(t), \quad \text{for} \quad t \in \mathcal{I}. \tag{5.53}$$

This proves that $|v_{1:2}| = |\tilde{v}_{1:2}|$ for $t \in \mathcal{I}$. Differentiating this equality, which is possible since $\mathcal{I}$ is a relatively open subset of $[0, t^*]$, and using again (5.41) allows to conclude the first step of the proof. □

Proof of Proposition 6.13. From (5.38) we have $\Gamma^0_0(v_0, 0) = \sigma(v_0)$. From Lemma 6.12 one has for all $\rho \geq 0$ that $\Gamma^1_1(v_0, \rho) \leq 0$ (respectively $\Gamma^1_1(v_0, \rho) \geq 0$) for all $v_0 \geq 0$ (respectively for all $v_0 \leq 0$). From the fact that $\partial_\rho \Gamma^0_0 = \Gamma^1_1$ and by continuity, this yields

$$|\Gamma^0_0(v_0, \rho)| \leq |\sigma(v_0)|, \quad \forall (v_0, \rho) \in \mathbb{R} \times [0, \infty). \tag{5.54}$$

From Assumption 6.1 on σ, one has $|\sigma(v_0)| \leq \sigma'(0)|v_0|$. Substituting this in (5.41), together with the lower bound on $I_0(t)$ yields

$$\dot{v}_0 \geq -|v_0| - |J_0|\sigma'(0)|v_0| + c, \tag{5.55}$$

whence, for any solution $v(\cdot)$ of (5.1) with an input satisfying the assumptions of the proposition,

$$\dot{v}_0(t) > 0 \quad \text{for all } t \text{ such that} \quad |v_0(t)| < \delta^* . \tag{5.56}$$

Similarly, the time estimation is obtained from applying Grönwall's lemma to (5.55).

Now, consider a solution $v(\cdot)$ and remember that $\delta > 0$ is smaller than δ^*. If $v_0(0) < -\delta$, either $v_0(0)$ remains smaller than $-\delta$ for all time, and we are in the first case of the proposition, or there is a smaller time $t_1 > 0$ such that $v_0(t_1) = -\delta$, and in that case, (5.56) implies that we must be in the second case of the proposition because $v_0(t)$ increases as long as it is smaller than δ^*, hence it passes through δ at some finite time t_2 and never becomes smaller than δ in the future. If $-\delta \leq v_0(0) < \delta$, by the same argument, there is a time t_2 so that $v_0(t)$ has the behavior describes as the third case of the Proposition, and if $v_0(0) \geq \delta$, which is the last possibility, the same arguments tell us that $v_0(t)$ is larger than δ for all positive t. □

5.5 THE OBSERVABILITY MAPPING

In this section we focus on the observability of system (5.1). For any solutions of (5.1) with measurement $h(v) = v_0$, we want to show the injectivity between the function $v_0(\cdot)$ and its time derivatives, up to certain order, with the solution $v(\cdot)$.

To do so, we look at the properties of the mapping T defined in (5.18), and in particular we aim at constructing the pseudo-inverse of Theorem 5.3.6. In order to do so, we start by extending T_t to a diffeomorphism, for each time $t \geq 0$.

5.5.1 *Observability mapping extension and differential observability*

From the symmetry of the system, we want to study T from a polar coordinate point of view. Letting $X = (v_0, \rho, e_\phi)^\top \in \mathbb{R} \times (0, \infty) \times \{\zeta \in \mathbb{R}^2 : |\zeta|^2 = 1\}$, (recall $v_{1:2} = \rho e_\phi$ with $e_\phi = (\cos\phi, \sin\phi)^\top$), we obtain from (5.41) the following dynamics in polar coordinates

$$\dot{X} = \begin{bmatrix} -v_0 + J_0 \Gamma_0^0(v_0, \rho) + I_0 \\ -\rho + J_1 \Gamma_0^1(v_0, \rho) + I_{1:2}^\top e_\phi \\ \frac{1}{\rho}(-(I_{1:2}^\top e_\phi) e_\phi + I_{1:2}) \end{bmatrix} . \tag{5.57}$$

We can naturally extend (5.57) by changing the variable $e_\phi \in S^1$, where S^1 is the unit circle of $\mathbb{R}^2$ to $\zeta \in \mathbb{R}^2$. We obtain then a new dynamic $\dot{X} = F(X, t)$ on $\mathbb{R} \times (0, +\infty) \times \mathbb{R}^2$ as follow

$$F(X,t) = \begin{bmatrix} -v_0 + J_0 \Gamma_0^0(v_0, \rho) + I_0 \\ -\rho + J_1 \Gamma_0^1(v_0, \rho) + I_{1:2}^\top \zeta \\ \frac{1}{\rho}(-(I_{1:2}^\top \zeta)\zeta + I_{1:2}) \end{bmatrix}, \qquad \text{where} \quad X = \begin{pmatrix} v_0 \\ \rho \\ \zeta \end{pmatrix} \in \mathbb{R} \times (0, +\infty) \times \mathbb{R}^2. \tag{5.58}$$

We define

$$\mathcal{X} := \mathbb{R}^* \times (0, \infty) \times \mathbb{R}^2. \tag{5.59}$$

Let us assume that $I \in C^2([0,+\infty),\mathbb{R}^3)$. With $\mathcal{L}$ the operator defined in (5.17), and omitting the arguments of Γ_p^j (defined in (5.38)), we define $S_t : \mathcal{X} \to \mathbb{R}^4$ by

$$\begin{aligned}
[S_t(X)]_0 &= v_0 \\
[S_t(X)]_1 &= -v_0 + J_0\Gamma_0^0 + I_0 \\
[S_t(X)]_2 &= -F_0 + J_0\Gamma_1^0 F_0 + J_0\Gamma_1^1 F_1 + \dot{I}_0 \\
[S_t(X)]_3 &= -\mathcal{L}F_0 + J_0\Gamma_1^0\mathcal{L}F_0 + J_0\Gamma_2^0(F_0)^2 \\
&\quad + 2J_0\Gamma_2^1 F_1 F_0 + J_0\Gamma_1^1\mathcal{L}F_1 + J_0\Gamma_2^2(F_1)^2 + \ddot{I}_0.
\end{aligned} \tag{5.60}$$

From direct computations, one can see that for any $t \geq 0$, any $X \in \mathbb{R}^* \times (0,+\infty) \times \{\zeta \in \mathbb{R}^2 : |\zeta| = 1\} \to \mathbb{R}^4$ it holds

$$S_t\,(X) := T_t(X_0, X_1 X_{2:3}). \tag{5.61}$$

Recall that by Assumption 5.3.5 we can fix an arbitrary large $R > 0$ such that $B_{\mathbb{R}^3}(0,R)$ is an invariant attracting set for (5.1). For any $\delta > 0$, $\eta > 0$, we set

$$\mathcal{X}_{\delta,\eta} = \{(v_0,\rho,\zeta) \in \mathcal{X} : \ \delta \leq |v_0| \leq R,\ \eta \leq \rho \leq R,\ |\zeta| \leq \ R\}. \tag{5.62}$$

We then have the following.

Theorem 5.5.1. *Assume $I \in C^2([0,+\infty),\mathbb{R}^3)$. Let $t \geq 0$ be such that $I_{1:2}(t) \wedge \dot{I}_{1:2}(t) \neq 0$. The map $S_t : \mathcal{X} \to \mathbb{R}^4$ is a C^∞ diffeomorphism onto its image. In particular, under Assumption 5.3.5 the map S_t is invertible for all times $t \geq 0$ and the inverse $S_t^{-1} : S_t(\mathcal{X}) \mapsto \mathcal{X}$ satisfies*

$$\sup_t \left\{ \|\partial_z S_t^{-1}(z)\| : z \in S_t(\mathcal{X}_{\delta,\eta}) \right\} < \infty. \tag{5.63}$$

Proof. Let $t \geq 0$ be such that $I_{1:2}(t) \wedge \dot{I}_{1:2}(t) \neq 0$. The map S_t is defined as composition of the functions Γ_p^j and $\mathcal{L}^i F$, $0 \leq i \leq 2$. From equation (5.58), F is defined for all $X \in \mathcal{X}$. As such, it is C^∞ over $\mathcal{X}$. In the following we denote elements of $\mathcal{X}$ by $X = (v_0,\rho,\zeta)$, where $v_0 \in \mathbb{R}^*$, $\rho \in (0,+\infty)$, and $\zeta \in \mathbb{R}^2$.

Step 1. Injectivity of S_t. Let $X, \tilde{X} \in \mathcal{X}$ be such that $S_t(X) = S_t(\tilde{X})$. We will compare term by term the expression of S_t given in (5.60). Observe that since $v_0 \neq 0$ by definition of $\mathcal{X}$, Lemma 6.12 implies injectivity of $\Gamma_0^0(v_0,\cdot)$ over $(0,\infty)$ and $\Gamma_1^1(v_0,\cdot) \neq 0$.

By the expression of the first two components of S_t, we immediately have

$$v_0 = \tilde{v}_0 \quad \text{and} \quad \rho = \tilde{\rho}.$$

Using the expression (5.58) of F one verifies that all terms in the expression of $[S_t]_2$ depend only on ρ and v_0, except $\Gamma_1^1(v_0,\rho)F_1(X,t)$. The fact that $\Gamma_1^1(v_0,\rho) \neq 0$, and the expression of $F_1(X,t)$ then implies that

$$I_{1:2}^\top \zeta = I_{1:2}^\top \tilde{\zeta}. \tag{5.64}$$

We now turn to the expression of $[S_t]_3$. By using again the expression (5.58) of F, eliminating all terms that depend only on ρ and v_0, and using the fact that $\Gamma_1^1(v_0,\rho) \neq 0$, we obtain

$\mathcal{L}F_1(X,t) = \mathcal{L}F_1(\tilde{X},t)$, which is the term that contains the last unknown term. Developing this expression and using (5.64), we obtain

$$\dot{I}_{1:2}^\top \zeta = \dot{I}_{1:2}^\top \tilde{\zeta}. \tag{5.65}$$

Putting together (5.64) and (5.65), we finally obtain

$$\begin{bmatrix} I_{1:2}^\top \\ \dot{I}_{1:2}^\top \end{bmatrix} [\zeta - \tilde{\zeta}] = 0.$$

This implies that $\zeta = \tilde{\zeta}$ since $I_{1:2}(t) \wedge \dot{I}_{1:2}(t) \neq 0$. This ends the proof of the injectivity.

Step 2. S_t is a diffeomorphism onto its image. As we have seen that S_t is differentiable and injective, by the rank theorem we just need to show that its differential $\partial_X S_t$ is of maximal rank on $\mathcal{X}$. For any $X = (v_0, \rho, \zeta) \in \mathcal{X}$, direct computations yield

$$\partial_X S_t(X) = \begin{bmatrix} 1 & 0 & 0_{1\times 2} \\ \star & J_0 \Gamma_1^1(v_0,\rho) & 0_{1\times 2} \\ \star & \star & J_0\Gamma_1^1(v_0,\rho) G(v_0,\rho,\zeta) \end{bmatrix}. \tag{5.66}$$

Here, we let $0_{1\times 2} = (0,0) \in \mathbb{R}^{1\times 2}$, and

$$G(v_0,\rho,\zeta) := \begin{bmatrix} I_{1:2}^\top \\ \dot{I}_{1:2}^\top \end{bmatrix} + \tilde{U}(v_0,\rho,\zeta) \begin{bmatrix} 0_{1\times 2} \\ I_{1:2}^\top \end{bmatrix} \tag{5.67}$$

for some function $\tilde{U}$. We omit the expressions noted as $(\star)$ in (5.66) as they do not intervene at this level.

As we are under (5.5) and $X \in \mathcal{X}$, we have $\Gamma_1^1(v_0,\rho) \neq 0$ by Lemma 6.12. Therefore as the Jacobian matrix of S is triangular by block, to prove that it is invertible we are left to prove the invertibility of $G(v_0,\rho,\zeta)$. This follows by assumption, since

$$\det G(v_0,\rho,\zeta) = \dot{I}_{1:2} \wedge I_{1:2} \neq 0 \tag{5.68}$$

Therefore S_t is an immersion in $\mathcal{X}$.

Step 3. Estimates on the inverse with respect to time. The inverse diffeomorphism S_t^{-1} defined on $S_t(\mathcal{X})$ is such that $\partial_z S_t^{-1}(z) = (\partial_X S_t(S_t^{-1}(z)))^{-1}$. Since the dependence on t of $\partial_X S_t$ is only due to the presence of $I_{1:2}$, $\dot{I}_{1:2}$, I_0 and $\dot{I}_0$, and recalling that $J_0\Gamma_1^1$ is non null, we rewrite it as follows (we omit the dependencies on (v_0,ρ,ζ)):

$$\partial_X S_t(X) = \begin{bmatrix} 1 & 0 & 0_{1\times 2} \\ \alpha & J_0\Gamma_1^1 & 0_{1\times 2} \\ J_0\Gamma_1^1 \beta(I) & J_0\Gamma_1^1\gamma(I) & J_0\Gamma_1^1 G(I) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0_{1\times 2} \\ \alpha & J_0\Gamma_1^1 & 0_{1\times 2} \\ 0 & 0 & J_0\Gamma_1^1 \mathrm{Id}_2 \end{bmatrix}}_{=:M} \underbrace{\begin{bmatrix} 1 & 0 & 0_{1\times 2} \\ 0 & 1 & 0_{1\times 2} \\ \beta(I) & \gamma(I) & G(I) \end{bmatrix}}_{=:N(I)}, \tag{5.69}$$

for some function α independent of the input I and functions β, γ depending on I and its first derivative. Here, we denoted by $G(I)$ the matrix in (5.67), in order to stress its dependency on I. This expression yields that

$$|\partial_z S_t^{-1}| \leq |N(I)^{-1}||M^{-1}|$$

Observe that M is continuous on $\mathcal{X}$ (albeit singular if $\rho v_0 = 0$), and thus $\|M^{-1}\|_{S_t(\mathcal{X}_{\delta,\eta})} \leq c$ for some constant $c > 0$ depending only on δ and η. Moreover,

$$N(I)^{-1} = \begin{bmatrix} 1 & 0 & 0_{1\times 2} \\ 0 & 1 & 0_{1\times 2} \\ -G^{-1}(I)\beta(I) & -G^{-1}(I)\gamma(I) & G^{-1}(I) \end{bmatrix}.$$

The lower bound on the determinant of G given by Assumption 5.3.5 together with the fact that I and $\dot{I}$ are uniformly bounded with respect to time allows then to conclude. □

We now prove Theorem 5.3.4 as a corollary of Theorem 5.5.1. Recall the notations $\mathcal{Z}_0 = \{v \in \mathbb{R}^3, v_0 = 0\}$ and $\mathcal{Z}_{1:2} = \{v \in \mathbb{R}^3, v_1 = v_2 = 0\}$.

Proof of Theorem 5.3.4. Let $v \in \mathbb{R}^3 \setminus \mathcal{Z}_0$. Suppose at first $\rho = |v_{1:2}| > 0$. We have by construction that $T_t(v) = S_t(X)$ where $X = (v_0, \rho, \frac{v_{1:2}}{\rho}) \in \mathcal{M} := \mathbb{R}^* \times (0, +\infty) \times \{|\zeta| = 1\}$. Since Theorem 5.5.1 guarantees that S_t is a diffeomorphism of $\mathcal{X}$ on $S_t(\mathcal{X})$, it is therefore an injective immersion when restricted to the sub-manifold of dimension 3 $\mathcal{M} \subset \mathcal{X}$. It follows that T_t is also an injective immersion for $v \in \mathbb{R}^3 \setminus (\mathcal{Z}_0 \cap \mathcal{Z}_1)$.

If $|v_{1:2}| = 0$, let $\tilde{v}$ be such that $T_t(v) = T_t(\tilde{v})$. By computing the first two elements of the expression of T_t, one deduces that $\Gamma_0^0(v_0, |v_{1:2}|) = \Gamma_0^0(v_0, |\tilde{v}_{1:2}|)$. We deduce from Lemma 6.12 that this function is injective, hence that $|v_{1:2}| = 0$. Thus, $v = \tilde{v}$ and T_t is indeed injective on $\mathbb{R}^3 \setminus \mathcal{Z}_0$. □

Remark 5.10. *Following the demonstration of injectivity (Step 1), the expression for the inverse mapping S_t^{-1} can be directly derived. The sole non-analytical component in this construction is the inverse of the injective real-valued function $\Gamma_0^0(y(t), \cdot)$ at each time instance t for non-null values of $y(t)$. See figure 5.3 for more details.*

We can now prove the observability of the system as stated in Proposition 5.3.3.

Proof of Proposition 5.3.3. We start by proving *(ii)* $\Rightarrow$ *(i)*. To this aim, let v and $\tilde{v}$ be two solutions of (5.1) such that $h(v(t)) = h(\tilde{v}(t))$ for all $t \in [0, t^*]$. By the definition of T_t in (5.18) we have that $T_t(v(t)) = T_t(\tilde{v}(t))$ for all $t \in [0, t^*]$. By assumption and by continuity of $t \mapsto I_{1:2}(t) \wedge \dot{I}_{1:2}(t)$ there exists an open interval $\mathcal{I} \subset [0, t^*]$ such that $I_{1:2}(t) \wedge \dot{I}_{1:2}(t) \neq 0$ for all $t \in \mathcal{I}$. By the fact that $I_0(t) \geq c$ and Proposition 6.13, it follows that there exists $t_0 \in \mathcal{I}$ such that $v_0(t_0) \neq 0$. Thus, by Theorem 5.3.4 it follows that T_{t_0} is injective. This yields $v(t_0) = \tilde{v}(t_0)$, which proves the claim by uniqueness of solutions of (5.1).

We now turn to an argument *(i)* $\Rightarrow$ *(ii)*. In this case, due to the assumption that $I_0(t) > c$ for all $t \geq 0$, we get from Proposition 6.13 that there exist at most a single time $t_c \geq 0$ where $v(t_c) = 0$. From Proposition 5.3.1 we have that $|v_{1:2}| = |\tilde{v}_{1:2}|$ and $I_{1:2}^\top v = I_{1:2}^\top \tilde{v}$ a.e. If v and $\tilde{v}$ are two distinct solutions of (5.1) such as $v_0 = \tilde{v}_0$ and $I_{1:2}(0)^\top (v_{1:2}(0) - \tilde{v}_{1:2}(0)) = 0$ then it is direct to see that

$$h(v(t)) = h(\tilde{v}(t)), \ \forall t \geq 0 \quad \& \quad v(t) \neq \tilde{v}(t), \ \forall t \geq 0, \tag{5.70}$$

which ends the proof. □

5.5.2 *The pseudo-inverse*

The goal of this section is to construct a pseudo-inverse of the observability mapping T_t, that is, to construct a Lipschitz map $\mathfrak{T}_t : \mathbb{R}^4 \to \mathbb{R}^3$ such that $\mathfrak{T}_t \circ T_t$ is the identity on (most of) $\mathbb{R}^3$. This map $\mathfrak{T}_t$ is a fundamental building block of our observer. Under the assumptions of Theorem 5.5.1, it is not difficult to provide an inverse of the map T_t since, letting

$$\Phi(X) := (X_0, X_1X_2, X_1X_3), \quad X \in \mathbb{R}^4 \tag{5.71}$$

it holds

$$\Phi \circ S_t^{-1} \circ T_t = \mathrm{Id} \quad \text{on } \mathbb{R}^3 \setminus \mathcal{Z}. \tag{5.72}$$

The above cannot be directly used to obtain a Lipschitz inverse since, as can be seen in the proof of Theorem 5.5.1, the Lipschitz constant of S_t^{-1} blows up when ρ or v_0 are small. However, S_t^{-1} is Lipschitz on $S_t(\mathcal{X}_{\delta,\eta})$, which suggests defining an inverse via (5.72) by inserting a projection on $S_t(\mathcal{X}_{\delta,\eta})$ in the definition. However, due to the complicated structure of $S_t(\mathcal{X}_{\delta,\eta})$, this is unpractical, and thus we resort to the following definition for the pseudo-inverse $\mathfrak{T}_t$:

$$\mathfrak{T}_t = \Phi \circ \mathrm{cut} \circ S_t^{-1} \circ \Pi_t. \tag{5.73}$$

Here, Π_t is a projection depending on the output $y(t) = v_0(t)$ and cut is an appropriate cut-off function, bounding the non-linear part in the observer and thus insuring that trajectories do not explode in finite time.

Namely, observe that a necessary condition for z to belong to $S_t(\mathcal{X}_{\delta,\eta})$ is that

$$|z_0| > \delta, \qquad \text{and} \qquad \begin{cases} J_0\Gamma_0^0(z_0, R) \le z_1 + z_0 - I_0 \le J_0\Gamma_0^0(z_0, \eta) & \text{if } z_0 > 0, \\ J_0\Gamma_0^0(z_0, \eta) \le z_1 + z_0 - I_0 \le J_0\Gamma_0^0(z_0, R) & \text{if } z_0 < 0. \end{cases} \tag{5.74}$$

Hence, $\Pi_t : \mathbb{R}^4 \to \mathbb{R}^4$ is defined as the projection on the above set. That is,

$$\Pi_t(z)_0 = \begin{cases} \min\{R, \max\{\delta, z_0\}\} & y(t) \ge 0, \\ \max\{-R, \min\{-\delta, z_0\}\} & y(t) < 0, \end{cases} \tag{5.75}$$

$$\Pi_t(z)_1 = \begin{cases} g_{\eta,R}(\Pi_t(z)_0, z_1), & y(t) \ge 0, \\ g_{R,\eta}(\Pi_t(z)_0, z_1), & y(t) < 0, \end{cases} \tag{5.76}$$

$$\Pi_t(z)_2 = z_2, \tag{5.77}$$

$$\Pi_t(z)_3 = z_3. \tag{5.78}$$

Here, we let the auxiliary function g to be defined by

$$g_{\rho_1,\rho_2}(z_0, z_1) = \begin{cases} J_0\Gamma_0^0(z_0, \rho_1) - z_0 + I_0 & \text{if } z_1 > J_0\Gamma_0^0(z_0, \rho_1) - z_0 + I_0, \\ J_0\Gamma_0^0(z_0, \rho_2) - z_0 + I_0 & \text{if } z_1 < J_0\Gamma_0^0(z_0, \rho_2) - z_0 + I_0, \\ z_1 & \text{otherwise.} \end{cases} \tag{5.79}$$

It is clear that Π_t is a Lipschitz function with Lipschitz constant uniformly bounded w.r.t. time.

The image of the projection Π_t fails to be in $S_t(\mathcal{X}_{\delta,\eta})$ due to not respecting the R upper bound in its last two components. Although we can still inverse S_t^{-1} on the image of Π_t, in order to avoid solutions of the observer (6.16) from blowing up in finite time, we add an additional cut-off. Let $p : \mathbb{R} \to [0,1]$ be a smooth function such that $p(x) = 1$ if $|x| \leq R-1$, $p(x) = 0$ if $|x| \geq R$ (recall R has been assumed to be arbitrarily large, and in particular larger than 3). Then

$$\mathrm{cut}(X) = (X_0, X_1, p(|X_{2:3}|)X_{2:3}). \tag{5.80}$$

We are now in a position to prove Theorem 5.3.6.

Proof of Theorem 5.3.6. For each $t \geq 0$, the function $\mathfrak{T}_t : \mathbb{R}^4 \to \mathbb{R}^3$ introduced in Equation (5.73) is well defined. Since $\Pi_t(\mathbb{R}^4) = S_t(\{X \in \mathbb{R}^4,\ |v_0| > \delta, \eta < \rho < R, \zeta \in \mathbb{R}^2\})$ where the inverse S_t^{-1} is well defined and C^∞ due to Theorem 5.5.1, it is a composition of continuous functions. The image of the map $\mathfrak{T}_t$ is

$$\mathfrak{T}_t(\mathbb{R}^4) = \Phi \circ \mathrm{cut} \circ S_t^{-1} \circ \Pi_t(\mathbb{R}^4) = \Phi(\mathcal{X}_{\delta,\eta}) \subset \{\delta \leq |v_0| \leq R\} \times B_{\mathbb{R}^2}(0, R^2). \tag{5.81}$$

Moreover $\mathfrak{T}_t$ is globally Lipschitz continuous. Indeed, it is a composition of locally Lipschitz functions. The map $S_t^{-1} \circ \Pi_t$ is not globally Lipschitz continous however (due to the last two components). By construction of the cut-off function, we have

$$\|\partial_z \,\mathrm{cut} \circ S_t^{-1}\|_{\Pi_t(\mathbb{R}^4)} = \|(\partial_X \,\mathrm{cut})(\partial_z S_t^{-1})\|_{S_t(\mathcal{X}_{\delta,\eta})} \leq \|\partial_X \,\mathrm{cut}\| \, \|\partial_z S_t^{-1}\|_{S_t(\mathcal{X}_{\delta,\eta})}. \tag{5.82}$$

Since Φ is smooth and thus locally Lipschitz continuous, and $\mathrm{cut} \circ S_t^{-1} \circ \Pi_t(\mathbb{R}^4) \subset \mathcal{X}_{\delta,\eta}$, we have by construction that $\mathfrak{T}_t$ is a globally Lipschitz continuous map for every time t. We also have that Π_t being globally Lpischitz continous uniformly with respect to time. Hence, as a consequence of (5.82) and Theorem 5.5.1, Lipschitz constant for $\mathfrak{T}_t$ exhibits uniform boundedness with respect to time.

We now provide an argument for (5.27). Let $z = T_t(v)$ for some $v \in \mathbb{R}^3$ such that v_0 has the same sign as the input $y(t)$ and $|v_0| > \delta$. If $\eta < |v_{1:2}| \leq R$, since $\Pi_t(z) = z$, $S_t^{-1}(z) = \left(v_0, |v_{1:2}|, \frac{v_{1:2}}{|v_{1:2}|}\right)$. It is then immediate to check that (5.73) reduces to (5.72), that is $\mathfrak{T}_t(z) = v$. On the other hand, if $0 < |v_{1:2}| \leq \eta$, direct computations show that

$$\mathfrak{T}_t(z) - v = \left(0, (\eta - |v_{1:2}|)\frac{v_{1:2}}{|v_{1:2}|}\right) \implies |\mathfrak{T}_t(z) - v| \leq \eta. \tag{5.83}$$

By continuity, the above inequality extends to $|v_{1:2}| = 0$. □

5.6 OBSERVER CONVERGENCE

In this section we prove Theorem 5.3.8 concerning the convergence properties of Observer (6.16).

We need the following Lemma, insuring that the non-linear part of the observer is globally Lipschitz.

Lemma 5.6.1. *Under Assumption 5.3.5, for every $t \geq 0$ the map $(\mathcal{L}^4 h)_t \circ \mathfrak{T}_t : \mathbb{R}^4 \to \mathbb{R}$ is Lipschitz continuous uniformly with respect to time. Here, we let $(\mathcal{L}^4 h)_t := (\mathcal{L}^4 h)(\cdot, t)$.*

Proof. Observe that $(\mathcal{L}^4 h)_t$ is well-defined for $t \geq 0$, since $I \in C^3([0,+\infty), \mathbb{R}^3)$ by Assumption 5.3.5. Moreover, $(\mathcal{L}^4 h)_t$ is locally Lipschitz, uniformly with respect to time. Indeed, being the composition of smooth functions, $(\mathcal{L}^4 h)_t$ is locally Lipschitz (actually smooth) for any $t \geq 0$. The uniformity with respect to time follows by observing that the time dependent terms in the Jacobian matrix $\partial_v (\mathcal{L}^4 h)_t$ are linear combinations of $I_i^{(j)}$, $i \in \{0,1,2\}$ and $j \in \{0,1,2,3\}$, which are bounded in time by Assumption 5.3.5.

Observe that by Theorem 5.3.6 the map $\mathfrak{T}_t$ is globally Lipschitz, uniformly with respect to time, and has bounded image. Therefore, $\mathcal{L}^4 h \circ \mathfrak{T}_t$ is also globally Lipschitz continuous uniformly with respect to time. □

We now prove the well definiteness of the observer (6.16).

Proof of Proposition 6.10. Thanks to Lemma 5.6.1 the first equation of (6.16) has globally defined solutions. The same is true for the second equation by Proposition 6.14. By Proposition 6.13 we have that there exists at most two times $t_2 > t_1 \geq 0$ such that $y(t_1) = -\delta$ and $y(t_2) = \delta$. Hence, $\hat{v}$ is piece-wise continuous with at most two jump discontinuities. However, $\hat{v}$ is continuous at t_1. Indeed, in this case the initial condition after the switch of dynamics is $\hat{v}(t_1^+) = \mathfrak{T}_t(\hat{z}(t_1^-))$. This completes the proof. □

Remark 5.11. *Let t_2 be the time such that $y(t_2) = \delta$. In this case, when switching dynamics in the observer (6.16) we have $\hat{z}(t_2^+) = T_t(\hat{v}(t_2^-))$. This can induce a discontinuity in $\hat{v}$ since*

$$\hat{v}(t_2^+) = \mathfrak{T}_t(\hat{z}(t_2^+)) = \mathfrak{T}_t \circ T_t(\hat{v}(t_2^-)). \tag{5.84}$$

Indeed, it could happen that $|\hat{v}_{1:2}(t_2^-)| < \eta$, and thus Theorem 5.3.6 cannot be used to guarantee that $\mathfrak{T}_t \circ T_t(\hat{v}(t_2^-)) = \hat{v}(t_2^-)$.

We are now in position to prove Theorem 5.3.8. This proof follows the conventional proof of the convergence of high gain design (see [111], [87] or [29] for example) with the necessary modifications to address the singularities associated with the set $\mathcal{Z}$.

Proof of Theorem 5.3.8. As $0 < \delta < \delta^*$ and $0 < \eta < R$, we proved in Proposition 6.10 that, under these assumptions, solutions of (6.16) are well defined globally.

We start by proving Point (i). We suppose at first that the output verifies $y(t_0) \geq \delta$. From Proposition 6.13 it follows that $y(t) \geq \delta$ for every time $t \geq 0$

Since we are assuming $y(t) \geq \delta$, by Theorem 5.3.6 and by the fact that $z = T_t(v)$ and $v_0(t) = y(t)$, we have

$$|v - \hat{v}| \leq |v - \mathfrak{T}_t(z)| + |\mathfrak{T}_t(z) - \mathfrak{T}_t(\hat{z})| \leq \eta\chi(v) + M_0|z - \hat{z}|, \tag{5.85}$$

where $M_0 > 0$ is a uniform bound of the Lipschitz constant of $\mathfrak{T}_t$, which depends on η and δ. We are thus left to estimate $|z - \hat{z}|$. We define $e := \hat{z} - z$ and consider for any $l > 0$ the following Lyapounov function candidate

$$\|e\|_l^2 := e^\top P_l e, \tag{5.86}$$

where P_l is the matrix with elements $P_l(i,j) := \frac{(-1)^{i+j}}{l^{i+j-1}} \cdot \frac{(i+j-2)!}{(i-1)!(j-1)!}$, for $i,j \in \{1,2,3,4\}$. From Lemma 5.6.1, we can set $L > 0$ to be an upper bound for the Lipschitz constant of $(\mathcal{L}^4 h)_t$ on the compact set $\mathcal{U} := \{(u_0,u_1,u_2) \in \mathbb{R}^3, \delta \leq |u_0| \leq R \text{ and } \eta \leq |u_{1:2}| \leq R^2\}$, which can be chosen to be uniform with respect to time $t \geq 0$.

Then, since $y = Cz$, it holds

$$\begin{aligned}\frac{1}{2}\frac{d\|e\|_l^2}{dt} &= e^\top P_l(A - P_l^{-1}C^\top C)e + e^\top P_l W\left((\mathcal{L}^4 h)(\mathfrak{T}_t(\hat{z})) - (\mathcal{L}^4 h)(v)\right),\\ &= -\frac{l}{2}\|e\|_l^2 + e^\top P_l W\left((\mathcal{L}^4 h)(\mathfrak{T}_t(\hat{z})) - (\mathcal{L}^4 h)(v)\right),\\ &\leq -\frac{l}{2}\|e\|_l^2 + \sqrt{\frac{\nu_{\max}}{\nu_{\min}}}LM_0\|e\|_l^2 + \chi(v)Ll^{-7/2}\eta\|e\|_l.\end{aligned} \tag{5.87}$$

Here, on the second line we used the fact that $P_l A + A^\top P_l - C^\top C = -lP_l$. The last line follows by Cauchy-Schwarz inequality and (5.85), thanks to the estimates

$$\|W\|_l \leq l^{-7/2}\sqrt{\nu_{\max}} \qquad \text{and} \qquad l^{-7}\nu_{\min}\,\mathrm{Id}_4 \leq P_l \leq l^{-1}\nu_{\max}\,\mathrm{Id}_4\,. \tag{5.88}$$

Here, $\nu_{\max}$ and $\nu_{\min}$ are respectively the max and the min eigenvalue of P_1.

We denote $l^* := 2\sqrt{\frac{\nu_{\max}}{\nu_{\min}}}LM_0$, and $M := M_0\frac{\max\{1,\sqrt{\nu_{\max}}\}}{\sqrt{\nu_{\min}}}$. Dividing both sides of (5.87) by $\|e\|_l$, we can apply a linear Grönwall's inequality to the variation of $\|e\|_l$. Since $\|e\|_l \geq \nu_{\min}l^{-7}|\hat{z} - z|$ and we have (5.85), we obtain the following estimation in the immersed domain, for $t \in [t_0, t_f]$,

$$|v(t) - \hat{v}(t)| \leq \eta\chi(v(t)) + Me^{-(l-l^*)t}\left(l^3|\hat{z}(t_0) - z(t_0)|e^{(l-l^*)t_0} + \int_{t_0}^{t} e^{(l-l^*)s}L\eta\chi(v(s))\,ds\right). \tag{5.89}$$

Using that $z(t) = T_t(v(t))$ for all times $t \geq t_0$ yields the first part of the point.

We now deal with the error increase during the switch. We recall that $t_1 := \inf\{t \geq 0 : |y(t)| < \delta\} = \inf\mathcal{S}$ and $t_2 = \sup\mathcal{S}$. We suppose now that $0 < t_1 < \infty$. We recall that by Proposition 6.13, it holds $t_2 \leq t_1 + t_\delta$, and $y(t) \geq \delta$ for all times $t \geq t_2$. If a switch occurs, the initial condition at t_1 for the second dynamics in (6.16) is $\hat{v}(t_1^+) = \mathfrak{T}_t(\hat{z}(t_1^-))$. The solution $\hat{v}$ is therefore well defined in the interval $[t_1, t_2]$ and continuous, since f is uniformly Lipschitz and $y(\cdot)$ is strictly increasing.

Let $L_1 > 0$ be the Lipschitz constant of T_t on the compact set $\mathcal{U}$, which is uniform with respect to time thanks to Assumption 5.3.5 (see, *e.g.*, proof of Lemma 5.6.1). Then, since $\hat{z}(t_2^-) = T_{t_2}(\hat{v}(t_2^-))$ and $z(t_2) = T_{t_2}(v(t_2))$, we have

$$|\hat{z}(t_2^-) - z(t_2^-)| \leq L_1|\hat{v}(t_2^-) - v(t_2^-)|. \tag{5.90}$$

Letting $L_2 > 0$ be the Lipschitz constant of $f(\cdot,t)$ on the compact $\mathcal{U}$, which is uniform with respect to time (see, *e.g.*, Proposition 6.14), we have

$$|\hat{v}(t_2^-) - v(t_2^-)| \leq e^{(t_2-t_1)L_2}|\hat{v}(t_1) - v(t_1)| \leq e^{t_\delta L_2}|\hat{v}(t_1) - v(t_1)|. \tag{5.91}$$

We have therefore characterised a bound of the error for all times $t \geq 0$ using both relations (5.89) and (5.91).

Let us establish Point (ii) of the theorem. Under the assumptions of the statement, $\chi(v(t)) = 0$ for all times $t \geq 0$. Recalling that $\mathcal{S} = [t_1, t_2]$, we will distinguish two cases: $t_1 = 0$ or $t_1 > 0$. In the first case, by the second part of Point (i), for any $t \leq t_2$ we have

$$|\hat{v}(t) - v(t)| \leq e^{L_2 t_\delta} |\hat{v}(0) - v(0)| \leq e^{(L_2 + (l-l^*))t_\delta} e^{-(l-l^*)t} |\hat{v}(0) - v(0)|. \tag{5.92}$$

Here we used that $e^{-(l-l^*)t_1} \leq e^{-(l-l^*)t} e^{(l-l^*)t_\delta}$ as $t \leq t_2$. Letting now $t \geq t_2$ and using the first part of Point (i) with $t_0 = t_2$ we get

$$|\hat{v}(t) - v(t)| \leq M l^3 e^{(L_2 + (l-l^*))t_\delta} e^{-(l-l^*)t} |\hat{v}(0) - v(0)|. \tag{5.93}$$

Hence, the statement stands proved in this case, with $k = M l^3 e^{(L_2 + (l-l^*))t_\delta}$. The argument for the case $t_1 > 0$ is similar, except that one has to start by applying the first part of Point (i) to $t \in [0, t_1]$, then plug this estimate in the error during the switching time for $t \in [t_1, t_2]$, and finally reapply the first part of Point (i) to obtain the estimate for $t \geq t_2$.

We finally provide an argument for Point (iii). We start by observing that, since $v(0) \in B_{\mathbb{R}^3}(0, R)$ and $\hat{v}$ is initialized in the compact set $\mathcal{K}$, it holds $|\hat{z}(0) - z(0)| \leq L_1(\operatorname{diam}(\mathcal{K}) + R)$. Then, one proceeds exactly as for the previous point except that instead of having $\chi(v(t)) = 0$ for all $t > 0$, one uses the trivial bound $\chi(v(t)) \leq 1$. This leads to the following estimates:

$$e^{-(l-l^*)t} \int_0^t e^{(l-l^*)s} \chi(v(s))\, ds \leq \frac{1}{l-l^*} \quad \text{and} \quad e^{-(l-l^*)t} \int_{t_2}^t e^{(l-l^*)s} \chi(v(s))\, ds \leq \frac{1}{l-l^*} \tag{5.94}$$

We assume δ to be sufficiently small, to be fixed later, so that $t_\delta < t^*$. In the case where the first switching time t_1 is positive, letting $C > 0$ be a constant independent of δ, η, and l one then obtains that for any $t \in [t^*, t_2] \cup (t_2 + t^*, +\infty)$ it holds

$$\frac{1}{C} |\hat{v}(t) - v(t)| \leq \eta \left(1 + M l^3 e^{L_2 t_\delta} e^{-(l-l^*)t^*} \left(1 + \frac{M}{l-l^*} \right) + \frac{M}{l-l^*} \right) + M^2 l^6 e^{(L_2 + (l-l^*))t_\delta} e^{-(l-l^*)t}. \tag{5.95}$$

Here, one has to be careful due to the fact that M depends on η, and in particular $M \to +\infty$ as $\eta \downarrow 0$. One starts by fixing $\eta < \varepsilon/(4C)$. Then, with η fixed, one chooses l sufficiently large so that

$$M^2 l^6 e^{-(l-l^*)t} \leq \frac{\varepsilon}{4C}. \tag{5.96}$$

Assuming that $l - l^* \geq 1$, this also implies that $M/(l - l^*) \leq 1$, so that (5.95) implies

$$|\hat{v}(t) - v(t)| \leq \frac{\varepsilon}{4} \left(2 + l^{-3} e^{L_2 t_\delta} \frac{\varepsilon}{2C} \right) + \frac{\varepsilon}{4} e^{(L_2 + (l-l^*))t_\delta}. \tag{5.97}$$

Then, assuming l is large enough so that $l^{-3}\varepsilon/2C < 1$, yields

$$|\hat{v}(t) - v(t)| \leq \varepsilon \left(\frac{1}{2} + \frac{e^{(L_2 + (l-l^*))t_\delta}}{2} \right). \tag{5.98}$$

Finally, since t_δ tends to 0 as δ tends to 0, we can fix $\delta > 0$ such that $e^{(L_2 + (l-l^*))t_\delta} \leq 1$. This completes the proof in this case.

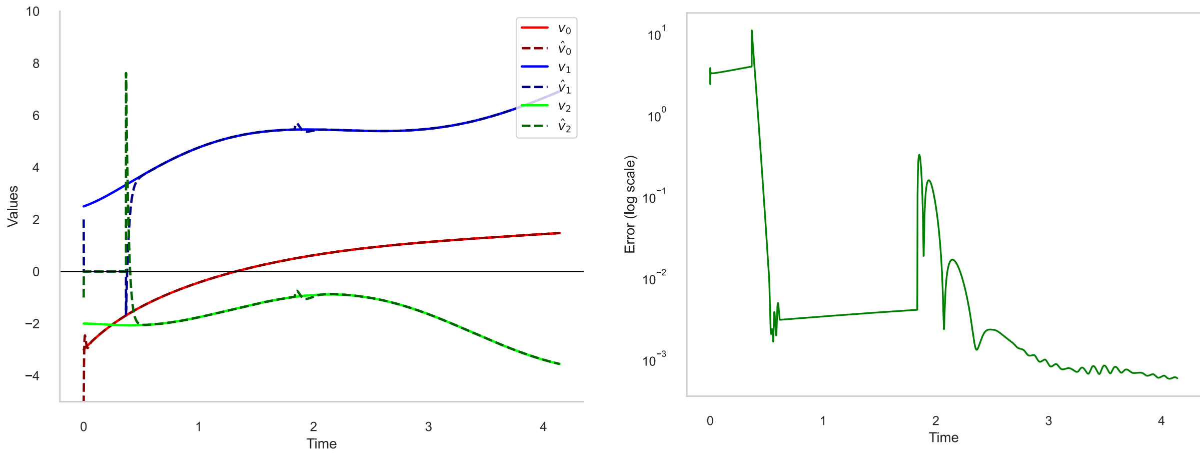


Figure 5.2: Left-hand side: components of the simulated trajectories of System (5.1) (plain lines) and Observer (6.16) (dashed lines). Right-hand side: plot of the estimation log-error. This particular solution satisfies point (ii) of the theorem. The switching time occurs at $t_1 \simeq 1.2$ and $t_2 \simeq 1.8$. Peaking phenomena occurs at the start of the simulation and at time t_2, when the high-gain observer restarts; they are better seen on the log-error plot. Since the error peaking at t_2 is proportional to the error at t_1, it is much smaller in size. At the beginning of the trajectory, $\hat{v}_1, \hat{v}_2$ are set to 0 due to the cut-off (5.80), until the observer $\hat{z}$ of the embedded system has sufficiently converged. This also results in a small increase in the error at around 0.5 second before a sharp decrease. At time t_1, the observer is turned off and the trajectories of v and $\hat{v}$ evolve according to the flow f of equation (5.1), resulting in a drift of the error over $[t_1, t_2]$.

The case where $t_1 = 0$ is treated similarly. Here, for any time $t < t_2$ the second part of Point (i) only yields a uniform constant bound. However, for any $t > t^* > t_2$, we obtain

$$\frac{1}{C}|\hat{v}(t) - v(t)| \leq \eta \left(1 + \frac{M}{l - l^*}\right) + Ml^3 e^{(L_2 + (l - l^*))t_\delta} e^{-(l - l^*)t}. \tag{5.99}$$

Here, $C > 0$ can be taken to be the same constant as in (5.95). It is then easy to see that the choice of η, l, and δ, considered in the case $t_1 > 0$ imply the statement also in this case. $\square$

5.7 NUMERICAL SIMULATIONS

We propose a numerical simulation for the observer (6.16). The parameters of the dynamical system (5.1) were chosen as such: $J_0 = -1$, $J_1 = 1.5$, $\sigma = \tanh(\mu \cdot -h_0)$ with the non linear gain $\mu = 10$ and the threshold $h_0 = 1$. The distribution P has been chosen as the Dirac mass $P = \delta_{r=1}$ for theoretical purposes. The input has been chosen as $I_0(t) = \varepsilon(1 - \beta)$, $I_{1:2}(t) = \beta\varepsilon\big(\cos(\frac{2\pi}{10}t), \sin(\frac{2\pi}{10}t)\big)$, with $\beta = 10^{-1}$ and $\varepsilon = 10^{-1}$. The initial condition was taken as $v(0) = (-3, 2.5, -2)$ so that the output verifies $y(0) < -\delta$ to show the switching mechanism. The timescale has been chosen as $\tau = 5$ for theoretical purposes to better show the error increase during the switch. The initial condition of the observer was taken as $\hat{v}(0) = (-5, 2, -1)$, $\hat{z}(0) = T_0(\hat{v}(0))$. As $J_0 < 0$, we chose $\delta = 3 * 10^{-1}$, which is greater than

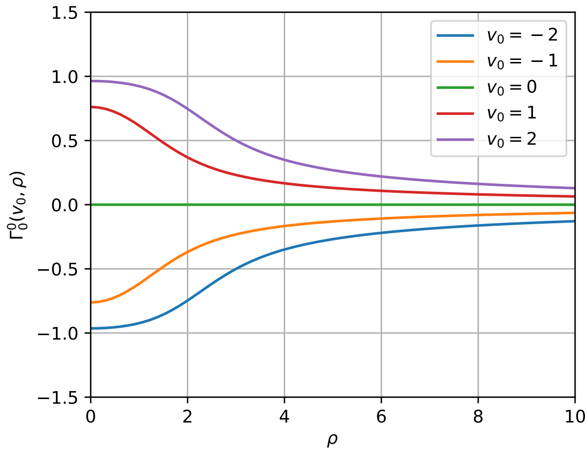


Figure 5.3: Plot of the function $\Gamma_0^0(v_0, \cdot)$ at different values of v_0. The figure showcases the observability singularity at $v_0 = 0$, as described in Proposition 5.3.3, and its injectivity for non-null values of v_0, as stated in point (*iii*) Lemma 6.12.

the robust bound of Proposition 6.13 (See the remark thereafter). The inversion error was taken $\eta = 10^{-3}$, the high gain $l = 15$. The numerical scheme chosen was an explicit RK4 with a step of 10^{-5}.

Concerning the computation of the pseudo-inverse $\mathfrak{T}_t$ defined in (5.73), this process relies on the inverse mapping S_t^{-1}. The expression for S_t^{-1} can be derived as outlined in Remark 5.10. Specifically, this requires the inversion of the function $\Gamma_0^0(y(t), \cdot)$ at each time step where $|y(t)| \geq \delta$, which was achieved using a Newton-based numerical algorithm. The resulting trajectory $\hat{v}$ is shown in Figure 5.2 for $t \in [0, 4]$. The function Γ_0^0 is showcased in figure 5.3.

APPENDIX: MODELISATION ADJUSTMENTS

Some authors favor a neural field model for activity instead of post-membrane potential, *e.g.*, [19, 33]. The model is similar to (6.1):

$$\tau \partial_t A(x,t) = -A(x,t) + \sigma\big(J \cdot A(\cdot,t) + I_{\text{ext}}(t)\big). \tag{5.100}$$

Supposing I_{ext} smooth enough, we can easily go from (5.100) to (6.1) by considering the change of variable: $V = J \cdot A + I_{\text{ext}}$, and replacing I_{ext} with $I_{\text{ext}} + \dot{I}_{\text{ext}}$. We favor (6.1) because it appears easier to study from a mathematical point of view. If our measurement is the averaged neural activity $a_0 = \int_\Omega A(u)du$, the change of variable $V = J \cdot A + I_{\text{ext}}$ leads to $\int_\Omega V(u)du = \langle A, J \cdot 1\rangle + \langle I_{\text{ext}}, 1\rangle = J_0 a_0 + I_0$. Therefore knowing the parameters J_0 and I_0, we can retrieve the mean membrane potential v_0 based on measurement not done in voltage but in spark per second (activity).

We also have some freedom in changing the shape of the nonlinearity. Some model choose sigmoidal functions $\sigma_+ : \mathbb{R} \to \mathbb{R}$ that are strictly positive, strictly increasing (i.e., $\sigma'_+ > 0$), convex-concave (i.e., $x\sigma''_+(x) \leq 0$ for all $x \in \mathbb{R}$), and such that

$$\lim_{x\to-\infty} \sigma_+(x) = 0, \quad \lim_{x\to+\infty} \sigma_+(x) = 1, \quad \sigma'_+ > 0, \quad \max_{\mathbb{R}} \sigma'_+ = \sigma'_+(0).$$

In this case there exist two parameters $s_1 > 0$ and $s_2 > 0$ such that $\sigma_+ = s_1\sigma + s_2$ where σ satisfies Assumption 6.1. Therefore, it suffices to replace I_0 by $I_0 + s_2$ and J_i by $J_i s_1$ for $i \in \{0, 1\}$.

On the other hand, if there is a threshold in the sigmoid function, then there exists a constant h_0 such that $\sigma(\cdot)$ is replaced by $\sigma(\cdot - h_0)$. Hence, we can always go back to our main equation by taking the change of variable from v_0 to $v_0 - h_0$. Replacing then I_0 with $I_0 + h_0$ maintains the performed analysis.

6

HOMOGENEOUS OBSERVER FOR A LOW-DIMENSIONAL NEURAL FIELDS MODEL OF CORTICAL ACTIVITY (PUBLISHED CDC 2024)

This chapter reproduces the article [**?**], published in *IEEE Conference on Decision and Control (CDC) 2024*.

We propose an observer design for a 3-dimensional model of cortical activity dynamics in the visual cortex, under the measurement of the averaged activity. It is based on the construction of an embedding of the system into a triangular 4-dimensional system where the dynamics are not Lipschitz-continuous but bear Hölder-continuity properties allowing us to implement a sliding mode observer. We discuss stability of the convergence of the observer with respect to relevant perturbations and provide simulations illustrating the method.

6.1 INTRODUCTION

The introduction of control theory techniques to the field of neurosciences is a rich and challenging issue. Its many applications range from brain health monitoring [149] to brain-computer interfacing [16, 150]. For instance, stabilization of signals in the brain via feedback control was used to alleviate Parkinson's disease symptoms [62]. Developing efficient and reliable strategies for online estimation of neural activity is fundamental to the implementation of such techniques. In the present paper, we focus on the question of observer design for a family of low dimensional neural fields. Neural fields equations model the evolution of the distribution of neuronal activity over a cortical domain. These dynamics are obtained by averaging the activity of large groups of neurons while maintaining spatial considerations in the interactions between neurons. Neural fields provide a powerful theoretical framework for studying brain activity (see [36] for a discussion of large scale models and [153] for a review on neural fields models in particular), and have proved their ability to replicate observed phenomena.

Neural fields models have a rich history of application to the study of the primary visual cortex V1 (see, e.g., [38] and references therein). In this paper, we are interested in a low-dimensional neural fields model of V1 discussed in [33], that we call Blumenfeld et al. model (see also [162]). The main feature of V1 that is captured by this model is the fact that neurons are more likely to be activated (or burst) when a stimulus arriving from the retina (through the LGN) is in accordance with some preset "preference" of each neuron. In particular, Blumenfeld et al. model focuses on orientation sensitivity, where edge orientation in the field of view induces neuron activation in V1 according to their prefered orientation. The favored orientation of a neuron is denoted by $\theta \in [-\pi/2, \pi/2)$, and their selectivity (how strong is that preference) with $r \in [0, \infty)$. This allows to label neurons in a given region of V1 according to the pair (θ, r), and to discard any other distinctive information. As a result, one can define the averaged post-membrane potential V for neurons with label (θ, r) (which we use as a substitute for neuronal activity), and describe its evolution with neural fields. Naturally, $V(\theta, r)$ must be π-periodic with respect to θ. Blumenfeld et al. model proposes to further reduce the dynamical model by focusing on the first few modes of the periodic function V, in order to detect the orientation corresponding to the highest activity in the considered region.

In that context, [8], by the authors of the present paper, discusses observability of the Blumenfeld et al. model from an control theory point of view. When the measurement is the post-membrane potential averaged over the whole region, as it happens with a non-localized electrode, we ask if it is possible to recover the most active orientation in the area. The analysis provided in [8] shows that this information is recoverable only under a persistent excitation condition. To test these conditions, the observability form was used to follow the standard high-gain observer design. This still required the introduction of a hybridization of the observer in order to cope with the observability singularities of the model, and limitations were imposed due to the Lipschitz-continuity failing in some regions. In the present paper, we discuss an observer design that is not based on the observability form, but rather on constructing an embedding of the dynamical model of Blumenfeld et al. that has triangular form and is such that the embedded dynamics are Hölder-continuous, allowing for the introduction of a sliding-mode observer design according to [26] (see also [118, 138]). Following these techniques yields convergence of the observer and also allows for a discussion of the stability of this convergence with respect to perturbation of the dynamics, that is extremely relevant in analyzing the behavior of this observer with respect to the reduction process on which the Blumenfeld et al. model is based.

6.2 MODEL DISCUSSION

Let us give an overview of the construction of the low-dimensional neural fields model from [33]. This discussion is developed further in [8]. We consider the evolution of the post-membrane potential V over an area of the cortex V1. Neurons of sensitivity $\theta \in [-\pi/2, \pi/2)$ and selectivity $r \in [0, +\infty)$ are grouped together. The distribution is assumed to be uniform in θ and to follow a (assumed to be known) probability distribution P in r. Without loss of generality for what follows, P is assumed to be compactly supported in $[0, +\infty)$. Then according to the measure $\mu = \frac{P(r)}{\pi} \mathrm{d}r\mathrm{d}\theta$ over $\Omega = [0, +\infty) \times [-\pi/2, \pi/2)$, we define V as an $L^2_\mu(\Omega)$ function evolving in time according to a single-layer neural field equation

$$\begin{cases} \tau \partial_t V(t) = -V(t) + J \cdot \sigma(V(t))(r, \theta) + I_{\text{ext}}(t) \\ V(0) \in L^2_\mu(\Omega). \end{cases} \tag{6.1}$$

Here, $I_{\text{ext}} \in L^\infty(\mathbb{R}, L^2_\mu(\Omega)) \cap C^0(\mathbb{R}, L^2_\mu(\Omega))$ denotes the external input, $\tau > 0$ the temporal synaptic constant. The operator J models the global effect of neurons in Ω on neurons at (r,θ); it is given by a $L^2_{\mu\otimes\mu}$ integral kernel j over the domain Ω^2: $J \cdot U(r,\theta) = \int_\Omega j(r,\theta,r',\theta')U(r',\theta')\frac{P(r)}{\pi}\mathrm{d}r\mathrm{d}\theta$ for any $U \in L^2_\mu(\Omega)$. The firing rate function $\sigma : \mathbb{R} \to \mathbb{R}$ models how a population "charges" or "discharges" and is typically chosen to be a smooth sigmoidal function. Equation (6.1) is well defined and has a unique global solution for each initial condition $V(0) \in L^2(\Omega)$ which has the same regularity as I_{ext}. See [38, 162] for instance.

We make the following standard assumptions on the sigmoid function.

Assumption 6.1. *The function σ is a $C^\infty(\mathbb{R},\mathbb{R})$ function that is odd, strictly increasing (i.e., $\sigma' > 0$), convex-concave (i.e., $x\sigma''(x) \leq 0$ for all $x \in \mathbb{R}$), and such that*

$$\lim_{x\to-\infty}\sigma(x) = -1, \quad \lim_{x\to+\infty}\sigma(x) = 1,$$
$$\sigma(0) = 0, \quad \max_{\mathbb{R}}\sigma' = \sigma'(0).$$

To arrive at the finite dimensional Blumenfeld et al. model, we make the reduction assumption that the interaction kernel j in operator J is of the form[1]

$$j(r,\theta,r',\theta') = J_0 + J_1 rr'\cos(2(\theta-\theta')), \quad J_1 > 0, J_0 \neq 0.$$

Then the set $L^2_\mu(\Omega)$ splits into $\mathrm{range}(J) + \mathrm{range}(J)^\perp$. The projection of V in $\mathrm{range}(J)$ takes the form $V^\parallel = v_0 + v_1 r\cos(2\theta) + v_2 r\sin(2\theta)$, and (6.1) yields (with $V^\perp = V - V^\parallel$)

$$\begin{cases} \tau\dot{V}^\parallel = -V^\parallel + J\cdot\sigma(V^\parallel + V^\perp) \ + I^\parallel, \\ \tau\dot{V}^\perp = -V^\perp + I^\perp. \end{cases} \tag{6.2}$$

Making the assumption that $I^\perp$ is negligible, $V^\perp$ converges exponentially towards 0. We henceforth assume $I^\perp = 0$, $V^\perp = 0$ and the dynamical system (6.1) is reduced to

$$\tau\dot{V}^\parallel = -V^\parallel + J\cdot\sigma(V^\parallel) \ + I^\parallel.$$

By definition of $V^\parallel$, this is in fact the evolution of the 3-dimensional state $(v_0, v_1, v_2) \in \mathbb{R}^3$, under a nonlinear dynamical system with linear input $(I_0, I_1, I_2) \in \mathbb{R}^3$ (following $V^\parallel$ still $I^\parallel = I_0 + I_1 r\cos(2\theta) + I_2 r\sin(2\theta)$). This dynamical model is the one we call Blumenfeld et al. model, originating from [33]. Note that this construction of the model is closer in presentation to the construction of the closely linked ring model of V1 (see, e.g., [162]).

Let us rewrite these dynamics in a manageable form. Without loss of generality, we now assume $\tau = 1$. We set (letting $\sigma^{(i)}$ be the i-th derivative of σ)

$$\Gamma_i^j(v_0,\rho) = \int_\Omega (r\cos 2\theta)^j \sigma^{(i)}(v_0 + \rho r\cos 2\theta)\frac{P(r)}{\pi}\mathrm{d}r\mathrm{d}\theta.$$

1 This can be understood as a truncation of a Fourier expansion in $\theta - \theta'$ of $j(r,\theta,r',\theta') = J_0 + rr' j_1(\theta-\theta')$ with even j_1.

With $\bar{v} = (v_1, v_2)$, $\bar{I} = (I_1, I_2)$ and $v = (v_0, \bar{v})$, $\bar{I} = (I_0, \bar{I})$, the Blumenfeld et al. model takes the form

$$\begin{cases} \dot{v}_0 = -v_0 + J_0 \Gamma_0^0(v_0, |v|) + I_0, \\ \dot{\bar{v}} = -\bar{v} + J_1 \Gamma_0^1(v_0, |\bar{v}|) \frac{\bar{v}}{|\bar{v}|} + \bar{I}. \end{cases} \tag{6.3}$$

This expression is singular at $|\bar{v}| = 0$ but the singularity can be checked to be resolvable (and the dynamics to be smooth) thanks to Lemma 6.12 in appendix. Assuming $I \in C^0([0, +\infty), \mathbb{R}^3)$, for any $v(0) \in \mathbb{R}^3$, solutions to the Cauchy problem are are well defined for any positive time. Furthermore, under the assumption that $\sup_{[0,\infty)} |I| < \infty$, there exist $R^* > 0$ such that for any $R \geq R^*$ the ball $B_{\mathbb{R}^3}(0, R)$ is invariant by the dynamics. See Proposition 6.14 in appendix.

Finally, regarding the observation of V, we discussed in introduction the problem of estimation of V under the knowledge of the averaged potential V, so that

$$y = h(V(t)) = \int_\Omega V(r, \theta, t) \frac{P(r)}{\pi} \mathrm{d}r \mathrm{d}\theta. \tag{6.4}$$

Under the assumption that $V(r, \theta, t) = v_0(t) + v_1(t) r \cos(2\theta) + v_2(t) r \sin(2\theta)$, we obtain the input-output formulation

$$\begin{cases} \dot{v}_0 = -v_0 + J_0 \Gamma_0^0(v_0, |v|) + I_0, \\ \dot{\bar{v}} = -\bar{v} + J_1 \Gamma_0^1(v_0, |\bar{v}|) \frac{\bar{v}}{|\bar{v}|} + I, \\ y = v_0. \end{cases} \tag{6.5}$$

An observability analysis of the system was carried in [8], let us recall the main result of that study.

Proposition 6.2. *Assume the regularity $I \in C^3([0, +\infty), \mathbb{R}^3)$ for the input.*

- *If there exists $t_0 \in \mathbb{R}_+$ such that $y(t_0) = 0$, then (6.5) is not differentiably observable at t_0. As a result, if $I_0 \equiv 0$ on an interval $[t_1, t_2] \subset \mathbb{R}_+$ then the system is not observable on that interval.*
- *Assume $I_0 \geq c > 0$. Then (6.5) is observable on $[t_1, t_2] \subset \mathbb{R}_+$ if and only if $\det(\bar{I}, \dot{\bar{I}}) \not\equiv 0$ on $[t_1, t_2]$.*

Recall differential observability is the injectivity of the map $v \mapsto (y, \dot{y}, \dots)$, see [87]. The C^3-regularity of the input is mostly necessary to be able to differentiate the output 3 times and we will drop this assumption in the rest of the paper. The observability singularity at $y = 0$ has to be dealt with if we want to consider trajectories that can have vanishing output. The assumption $I_0 \geq c > 0$, motivated by the neuroscience modeling, ensures at most one crossing of the singularity $y = 0$. Our goal over the following sections is to describe an observer design for this system based on an embedding that puts (6.5) in a triangular form.

6.3 EMBEDDING AND TRIANGULAR FORM

For the remainder of the paper, we make the assumption that $I \in C^1([0,+\infty), \mathbb{R}^3)$, at least. Consider the mapping $\bar{G}_t : \mathbb{R}^4 \to \mathbb{R}^4$ defined for all $(v_0, \rho, \zeta) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2$ by

$$\bar{G}_t(v_0,\rho,\zeta) = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} v_0 \\ J_0\Gamma_0^0(v_0,\rho) \\ \langle \bar{I}, \zeta\rangle J_0\Gamma_1^1(v_0,\rho) \\ \langle \dot{\bar{I}}, \zeta\rangle J_0\Gamma_1^1(v_0,\rho) \end{pmatrix} \tag{6.6}$$

and $\Psi : \mathbb{R} \times (\mathbb{R}^2 \setminus \{0\}) \to \mathbb{R}^4$ defined by $\Psi(v_0, v) = (v_0, |v|, v/|v|)$. The embedding of (6.5) we wish to use is given by $G_t = \bar{G}_t \circ \Psi : \mathbb{R} \times (\mathbb{R}^2 \setminus \{0\}) \to \mathbb{R}^4$.

Lemma 6.3. *Assume $|\det(I, \dot{I})| > 0$. For all $t > 0$, the map G_t is an injective immersion in restriction to $\mathbb{R}^* \times (\mathbb{R}^2 \setminus \{0\})$. Furthermore, G_t can be continuously extended to $\mathbb{R}^3$ while retaining injectivity on $\mathbb{R}^* \times \mathbb{R}^2$.*

Proof. Injectivity of G_t is straightforward: assume $G(v) = G(\tilde{v})$, then $v_0 = \tilde{v}_0$. Denote $\rho = |\bar{v}|$, $\tilde{\rho} = |\tilde{\bar{v}}|$. From Lemma 6.12 in appendix, since $v_0 \neq 0$, $\rho = \tilde{\rho}$. Then since $\det(\bar{I}, \dot{\bar{I}}) \neq 0$ and $\Gamma_1^1(v_0,\rho) \neq 0$ (from Lemma 6.12 again), we get $v = \tilde{v}$.

Regarding the Jacobian matrix of G_t, we focus on the form $\bar{G}_t \circ \Psi$. The Jacobian matrix of Ψ has rank 3 on $\mathbb{R} \times (\mathbb{R}^2 \setminus \{0\})$ (it has 3×3 minors $-v_2/\rho^2$ and v_1/ρ^2 that cannot be both zero). On the other hand, the Jacobian matrix of $\bar{G}_t$ is block-lower-triangular, of the form $\operatorname{Jac} G_t(v_0,\rho,\zeta) = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}$ with

$$A_{11} = \begin{pmatrix} 1 & 0 \\ * & J_0\Gamma_1^1(v_0,\rho) \end{pmatrix}, A_{22} = \frac{J_0\Gamma_1^1(v_0,\rho)}{\rho} \begin{pmatrix} I_1 & I_2 \\ \dot{I}_1 & \dot{I}_2 \end{pmatrix}.$$

Since $v_0 \neq 0$, $\rho \neq 0$, $\det(\bar{I}, \dot{\bar{I}}) \neq 0$, the Jacobian matrix of $\bar{G}_t$ is invertible, and the Jacobian matrix of G_t necessarily achieves the necessary maximal rank 3 implying immersion.

Extension of the domain of G_t to $\mathbb{R}^* \times \mathbb{R}^2$ comes from the fact that as $\rho \to 0$, for any $v_0 \neq 0$, $\Gamma_1^1(v_0,\rho) \sim \Gamma_2^2(v_0,0)\rho$ ($\Gamma_2^2(v_0,0) \neq 0$ if and only if $v_0 \neq 0$). Hence G_t can be continuously extended at $(v_0, 0)$, for any v_0, by $G_t(v_0,0) = (v_0, \Gamma_0^0(v_0,0), 0, 0)$. $\square$

In order to gain uniformity with respect to time of the modulus of continuity of a potential left inverse of G_t, we assume the following.

Assumption 6.4. *The input $I(t) \in C^2([0,+\infty), \mathbb{R}^3)$ is bounded on $[0,+\infty)$, and so is its first derivative. Furthermore, the following lower bounds hold for some positive numbers $c, \mu > 0$:*

$$I_0(t) \geq c \text{ and } |\det(\bar{I}(t), \dot{\bar{I}}(t))| \geq \mu, \quad \forall t \geq 0. \tag{6.7}$$

Set $\mathcal{D} = \{(x_0, x_1) \mid x_0 \neq 0, x_1 \in J_0\Gamma_0^0(x_0, \mathbb{R})\}$ and for $0 < \delta < R$, set

$$\mathcal{D}_{\delta,R} = \{(x_0, x_1) \mid |x_0| \in [\delta, R], x_1 \in J_0\Gamma_0^0(x_0, [-R, R])\}.$$

Relying on the previous lemma, we define for any given output y of (6.5) a pseudo inverse of G_t that extends the domain of G_t^{-1} beyond the image of G_t. First, we define a projection $\Pi : \mathbb{R}^4 \to \mathcal{D}_{\delta,R} \times \mathbb{R}^2$ that we assume to implicitly depend on the output y in the following way:

- $\Pi_2(x) = x_2, \Pi_3(x) = x_3$.
- If $y \geq 0$, then $\Pi_0(x)$ is the projection[2] of x_0 on $[\delta, R]$, and $\Pi_1(x)$ is the projection of x_1 on $[J_0\Gamma_0^0(\Pi_0(x), R), J_0\sigma(\Pi_0(x))]$.
- If $y < 0$, then take $\Pi_0(x)$ as the projection of x_0 on $[-R, -\delta]$, while $\Pi_1(x)$ is the projection of x_1 on $[J_0\sigma(\Pi_0(x)), J_0\Gamma_0^0(\Pi_0(x), R)]$.

Then for any $x \in \mathbb{R}^4$, letting ρ be uniquely defined by $\Gamma_0^0(\Pi_0(x), \rho) = \Pi_1(x)/J_0$, we set $\mathcal{G}_t : \mathbb{R}^4 \times \mathbb{R} \to \mathbb{R}^3$

$$\mathcal{G}_t(x) := \left(\Pi_0(x), \frac{\rho}{J_0\Gamma_1^1(\Pi_0(x), \rho)} \begin{pmatrix} I_1 & I_2 \\ \dot{I}_1 & \dot{I}_2 \end{pmatrix}^{-1} \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} \right) \tag{6.8}$$

Remark 6.5. *The choice of having $\mathcal{G}_t$ depend on the sign of the output y is guided by the loss of injectivity of G_t in the region $\{v \in \mathbb{R}^3 \mid v_0 = 0\}$ (indicating loss of observability). This basically amounts to constructing a pseudo inverse on each of the two regions $x_0 \geq 0$ and $x_0 < 0$ and picking the one most adapted to the measured output.*

The pseudo inverse has been constructed to verify the following: for all $v \in \mathbb{R}^3$ such that $|v_0| \geq \delta, |v| < R$, and $v_0 y(t) > 0$,

$$\mathcal{G}_t(G_t(v)) = v. \tag{6.9}$$

We can prove that this extension of G_t^{-1} is uniformly continuous.

Lemma 6.6. *Let $0 < \delta < R$. Under Assumption 6.4, there exist a class $\mathcal{K}$ function ω such that for every time $t \geq 0$, for every $(x, \tilde{x}) \in \mathbb{R}^4 \times \mathbb{R}^4$,*

$$|\mathcal{G}_t(x) - \mathcal{G}_t(\tilde{x})| \leq \omega(|x - \tilde{x}|). \tag{6.10}$$

Proof. The introduction of map Π ensures that $\Pi(x) \in \mathcal{D}_{\delta,R}$, and that we can recover ρ from $\Pi(x)$. The maps Π_0, Π_1 being projections, they are 1-Lipschitz continuous. Furthermore, the function $\mathcal{G}_t$ is smooth in ρ: the singularity $\rho/\Gamma_1^1(\Pi_0(x), \rho)$ resolves again as $\Gamma_1^1(v_0, \rho) \sim \rho\Gamma_2^2(v_0, 0)$. Hence $m : (x_0, x_1) \mapsto \frac{\rho}{J_0\Gamma_1^1(\Pi_0(x), \rho)}$ is uniformly continuous and there exists ω', a class $\mathcal{K}$ function such that $|m(x_0, x_1) - m(\tilde{x}_0, \tilde{x}_1)| \leq \omega'(|x - \tilde{x}|)$. By Assumption 6.4, $a := \sup_{t \geq 0} \left\| \begin{pmatrix} I_1 & I_2 \\ \dot{I}_1 & \dot{I}_2 \end{pmatrix}^{-1} \right\| < \infty$. This implies that $|\mathcal{G}_t(x) - \mathcal{G}_t(\tilde{x})| \leq (1 + a)|x - \tilde{x}| + \omega'(|x - \tilde{x}|) =: \omega(|x - \tilde{x}|)$. □

Now we consider an embedding system (6.5) (see, for instance, [29, Chapter 4]).

Proposition 6.7. *Suppose Assumption 6.4 holds. There exists continuous functions $\phi_0 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $\phi_1 : \mathcal{D} \times \mathbb{R} \to \mathbb{R}$, and $\phi_i : \mathcal{D} \times \mathbb{R}^{i-1} \times \mathbb{R} \to \mathbb{R}$, $i = 2, 3$, such that solutions of (6.5) have their image by G_t be solutions of*

$$\begin{aligned} \dot{x}_0 &= \phi_0(x_0, t) + x_1 \\ \dot{x}_1 &= \phi_1(x_0, x_1, t) + x_2 \\ \dot{x}_2 &= \phi_2(x_0, x_1, x_2, t) + x_3 \\ \dot{x}_3 &= \phi_3(x_0, x_1, x_2, x_3, t). \end{aligned} \tag{6.11}$$

2 By projection of a on $[b, c]$ for three real numbers a, b, c, with $b \leq c$, we mean $\min(\max(a, b), c)$.

Proof. Let v be a solution of (6.5), and let $\rho = |\bar{v}|$, $\zeta = \bar{v}/|\bar{v}|$. Two useful formulas (well-posed when $\rho \neq 0$) are $\dot{\rho} = -\rho + J_1\Gamma_0^1(v_0,\rho) + \langle \bar{I}, \zeta\rangle$ and $\rho\dot{\zeta} = \bar{I} - \langle \bar{I}, \zeta\rangle\zeta$. Let us compute the ϕ_i functions. First,

$$\dot{x}_0 = -x_0 + x_1 + I_0 = \phi_0(x_0,t) + x_1 \tag{6.12}$$

The function ϕ_0 is continuous, naturally. Then

$$\begin{aligned}\dot{x}_1 =& J_0\Gamma_1^0(x_0,\rho)\dot{x}_0 + J_0\Gamma_1^1(x_0,\rho)\dot{\rho}\\ =& J_0\Gamma_1^0(x_0,\rho)(\phi_1(x_0)+x_1) - \rho J_0\Gamma_1^1(x_0,\rho)\\ &+ J_1J_0\Gamma_1^1(x_0,\rho)\Gamma_0^1(x_0,\rho) + J_0\Gamma_1^1(x_0,\rho)\langle \bar{I},\zeta\rangle\\ =& \phi_1(x_0,x_1,t) + x_2.\end{aligned} \tag{6.13}$$

The heart of the continuity argument is that the mapping $(v_0,\rho) \mapsto \Gamma_0^0(v_0,\rho)$ admits a smooth inverse on its image, as long as $v_0 \neq 0$ (see Lemma 6.12). Hence $(x_0,x_1) \mapsto (x_0,\rho)$ is continuous on $\mathcal{D}$, and so is ϕ_1 on its domain.

Regarding (x_2,x_3),

$$\begin{aligned}\dot{x}_2 =& \langle \dot{\bar{I}},\zeta\rangle J_0\Gamma_1^1(x_0,\rho) + \langle \bar{I},\dot{\zeta}\rangle J_0\Gamma_1^1(x_0,\rho)\\ &+ \langle \bar{I},\zeta\rangle J_0(\Gamma_2^1(x_0,\rho)\dot{x}_0 + \Gamma_2^2(x_0,\rho)\dot{\rho})\\ =& x_3 + \langle \bar{I},\dot{\zeta}\rangle J_0\Gamma_1^1(x_0,\rho)\\ &+ \langle \bar{I},\zeta\rangle J_0(\Gamma_2^1(x_0,\rho)\dot{x}_0 + \Gamma_2^2(x_0,\rho)\dot{\rho})\end{aligned} \tag{6.14}$$

and

$$\begin{aligned}\dot{x}_3 =& \langle \ddot{\bar{I}},\zeta\rangle J_0\Gamma_1^1(x_0,\rho) + \langle \dot{\bar{I}},\dot{\zeta}\rangle J_0\Gamma_1^1(x_0,\rho)\\ &+ \langle \dot{\bar{I}},\zeta\rangle J_0(\Gamma_2^1(x_0,\rho)\dot{x}_0 + \Gamma_2^2(x_0,\rho)\dot{\rho}).\end{aligned} \tag{6.15}$$

Assuming $\det(\bar{I},\dot{\bar{I}}) \neq 0$, we get $\langle \ddot{\bar{I}},\zeta\rangle J_0\Gamma_1^1(v_0,\rho) = \alpha x_2 + \beta x_3$ with $\ddot{\bar{I}} = \alpha\bar{I} + \beta\dot{\bar{I}}$. Since I has been assumed to have C^2-regularity, α,β are continuous. The remaining question is continuity of the other terms in the expressions of ϕ_2,ϕ_3 induced by (6.14)-(6.15). We prove it in the case of ϕ_2, the proof is the same for ϕ_3.

Taking (6.14), with the expressions for $\dot{\rho}$ and $\dot{\zeta}$, and replacing every $\langle \bar{I},\zeta\rangle$ with $x_2/J_0\Gamma_1^1(v_0,\rho)$, yields (omitting the variables in the Γ_i^j):

$$\begin{aligned}\phi_2(x_0,x_1,x_2,t) :=& \frac{\Gamma_1^1}{\rho}\|\bar{I}\|^2 - \frac{x_2^2}{\rho J_0\Gamma_1^1} + x_2\frac{\Gamma_2^1}{\Gamma_1^1}(x_1 - x_0 + I_0)\\ &+ \frac{x_2\Gamma_2^2}{\Gamma_1^1}\left(-\rho + \Gamma_1^1 + \frac{x_2}{J_0\Gamma_1^1}\right)\end{aligned}$$

As long as $\rho \neq 0$ (or equivalently $x_1 \neq J_0\Gamma_0(v_0,0)$), every element of $\phi_2(x_0,x_1,x_2,t)$ is a smooth function of (x_0,x_1,x_2) ($\Gamma_1^1(v_0,\rho) = 0$ if and only if $\rho = 0$). Hence the only difficulty is showing that $\phi_2(x_0,x_1,x_2)$ has a well defined limit when $\rho \to 0$ (or equivalently $x_1 \to J_0\Gamma_0(v_0,0)$). Using the fact that the maps Γ_i^j are smooth and satisfy $\Gamma_i^j(v_0,0) = 0$ when j is odd, it is clear that the only non-immediate indeterminate form is

$$\lim_{\rho\to 0}\frac{x_2^2}{J_0\Gamma_1^1}\left(\frac{\Gamma_2^2}{\Gamma_1^1} - \frac{1}{\rho}\right)$$

Since we have the Taylor expansions $\Gamma_1^1(x_0,\rho) = \Gamma_2^2(x_0,0) + \rho^2\Gamma_4^4(x_0,0)/2 + o(\rho^3)$, $\Gamma_1^1(x_0,\rho) = \rho\Gamma_2^2(x_0,0) + \rho^3\Gamma_4^4(x_0,0)/6 + o(\rho^4)$, we deduce that $\frac{\Gamma_2^2-\rho\Gamma_1^1}{\rho\Gamma_1^1} = \frac{\Gamma_4^4(x_0,0)}{3\Gamma_2^2(x_0,0)}\rho + o(\rho^2)$. This resolves the indeterminate form

$$\frac{x_2^2}{J_0\Gamma_1^1}\left(\frac{\Gamma_2^2}{\Gamma_1^1} - \frac{1}{\rho}\right) = \frac{x_2^2\Gamma_4^4(v_0,0)}{J_0 3\Gamma_2^2(x_0,0)^2} + o(\rho).$$

and proves continuity of ϕ_2. □

System (6.5) has then been embedded in the continuous system (6.11) that is of *triangular form*. In order to implement an observer design that is adapted to this triangular form, we check that the functions ϕ_i satisfy appropriate regularity.

Proposition 6.8. *Suppose Assumption 6.4 holds. Let $0 < \delta < R$ and let K be a compact subset of $\mathcal{D}_{\delta,R} \times \mathbb{R}^2$. With respect to $x \in K$ and uniformly in time, the functions ϕ_0, ϕ_1 are Lipschitz-continuous while ϕ_2, ϕ_3 are $\frac{1}{2}$-Hölder-continuous. As a result they satisfy the homogeneity conditions required by Table 6.1.*

Proof. Recall that if two real valued functions f, g are α-Hölder-continuous over the compact K ($f, g \in C^{0,\alpha}(K,\mathbb{R})$) then their product fg is also α-Hölder-continuous. The same goes for f/g if $g \neq 0$ on K. As can be seen from the expressions of ϕ_i, $i = 1,2,3$ (the case $i = 0$ doesn't warrant discussion), the only issue in terms of regularity comes from the functions Γ_i^j. The mapping $(v_0,\rho) \mapsto \Gamma_i^j(v_0,\rho)$ is smooth for any pair (i,j) but the (inverse) mapping $(x_0,x_1) \mapsto (x_0,\rho)$ is not. Hence ϕ_i being smooth in ρ (at $\rho = 0$ included) is not sufficient to claim that the maps may be Hölder-continuous. For this, we prove that the inverse of $(x_0,\rho) \mapsto (x_0, J_0\Gamma_0^0(x_0,\rho))$ is locally $\frac{1}{2}$-Hölder-continuous.

Since $\Gamma_0^0(x_0,\cdot)$ is an even smooth function, we can define the odd smooth function

$$\gamma(x_0,s) := \begin{cases} \Gamma_0^0(x_0,\sqrt{s}) - \Gamma_0^0(x_0,0) & s \geq 0 \\ \Gamma_0^0(x_0,0) - \Gamma_0^0(x_0,\sqrt{|s|}) & s < 0. \end{cases}$$

For any $\rho \geq 0$, the relation $x_1 - \Gamma_0^0(x_0,\rho) = 0$ can be extended to $m(x_0,x_1,s) = x_1 - \gamma(x_0,s) = 0$ for any $x_0, x_1 \in \mathcal{D}$, $s \in \mathbb{R}$. Then $\partial_s m = \Gamma_2^2(x_0,\sqrt{|s|}) > 0$ on $\mathcal{D} \times \mathbb{R}$. This allows to use the implicit function theorem to define $s(x_0,x_1)$ and to check that the Jacobian matrix $\|\operatorname{Jac} s(x_0,x_1)\| \leq \frac{c}{\Gamma_2^2(|x_0|,\sqrt{|s(x_0,x_1)|})}$. Assuming $x \in K$, Γ_2^2 ends up lower bounded. Hence, applying the mean value theorem to γ, for any pair (x_0,x_1), $(\tilde{x}_0,\tilde{x}_1)$ and associated $\rho, \tilde{\rho}$, one has $|\rho^2 - \tilde{\rho}^2|^2 \leq C(|x_0 - \tilde{x}_0|^2 + |x_1 - \tilde{x}_1|^2)$. Since $|\rho - \tilde{\rho}|^2 \leq |\rho^2 - \tilde{\rho}^2|$ (as they're both non-negative), this concludes the Hölder-continuity proof. From this we deduce that ϕ_i, $i = 2,3$ are all $\frac{1}{2}$-Hölder-continuous as functions of (x_0,x_1,x_2,x_3) in restriction to the compact K, uniformity in time is immediate from Assumption 6.4.

What remains to be shown is that ϕ_1 is also Lipschitz-continuous. This is straight-forward from the definition of ϕ_1 found in (6.13). From this equation we deduce that ϕ_1 can be seen as a smooth function of ρ^2, which we have shown above to be a Lipschitz-continuous function of (x_0,x_1), hence the result. □

6.4 OBSERVER DESIGN

In this section, we propose an observer for system (6.5) and prove its main properties.

Let v be a solution of (6.5) with output y. We consider the following hybrid observer (which full description necessitates the discussion that follows its introduction).

$$\begin{cases} \text{If } |y(t)| > \delta : \\ \hat{v}(t) = \mathcal{G}_t(\hat{x}(t)), \quad \begin{cases} \dot{\hat{x}}_0 = \hat{\phi}_0 + \hat{x}_1 - Lk_0 \left[\hat{x}_0 - y\right]^{\frac{3}{4}} \\ \dot{\hat{x}}_1 = \hat{\phi}_1 + \hat{x}_2 - L^2 k_1 \left[\hat{x}_0 - y\right]^{\frac{1}{2}} \\ \dot{\hat{x}}_2 = \hat{\phi}_2 + \hat{z}_3 - L^3 k_2 \left[\hat{x}_0 - y\right]^{\frac{1}{4}} \\ \dot{\hat{x}}_3 \in \hat{\phi}_3 - L^4 k_3 \text{sign}(\hat{x}_0 - y). \end{cases} \\ \text{If } |y(t)| < \delta : \\ \dot{\hat{v}}(t) = f(\hat{v}, t), \qquad \hat{z}(t) = G_t(\hat{v}(t)), \end{cases} \tag{6.16}$$

where:

- f indicates the right-hand side of (6.5).
- L, k_i and δ are parameters to be tuned.
- For $z \in \mathbb{R}$, and $\alpha > 0$, $[z]^\alpha = \varepsilon|z|^\alpha$ where $\varepsilon = 1$ if $z > 0$, $\varepsilon = -1$ if $z < 0$ and $\varepsilon = 0$ if $z = 0$; and
$$\text{sign}(z) := \begin{cases} \{1\} & \text{if } z > 0, \\ [-1, 1] & \text{if } z = 0, \\ \{-1\} & \text{if } z < 0. \end{cases}$$
- With $M = 2|J_0| \sup_{\mathcal{D}_{\delta,R}} |\Gamma_1^1|$, $M_2 = M \sup_{\mathbb{R}_+} \|\bar{I}\|$ and $M_3 = M \sup_{\mathbb{R}_+} \|\dot{\bar{I}}\|$, we set $K = \mathcal{D}_{\delta,R} \times [-M_2, M_2] \times [-M_3, M_3]$
- The functions $\hat{\phi}_i$ are saturated versions of the functions ϕ_i. For $i = 0, 1, 2, 3$, take $S_i(z)$ to be the projection of $z \in \mathbb{R}$ onto $[-\sup_{x \in K} |\phi_i(x)|, \sup_{x \in K} |\phi_i(x)|]$, take $d_i(x) = (x_0, \ldots, x_i)$ and $\hat{\phi}_i = S_i \circ \phi_i \circ d_i \circ \Pi(x)$.

Remark 6.9. *The projection Π is added to ensure that the non-linearities ϕ_i are restricted to the domain $\mathcal{D}_{\delta,R} \times \mathbb{R}^2$ where they are well defined and verify the Hölder-conditions of Table 6.1 on any sub-compact. The saturations S_i are added due to the fact that $B_{\mathbb{R}^3}(0, R)$ is an invariant attracting ball if $R > 0$ is sufficiently large and that $G_t(B_{\mathbb{R}^3}(0, R)) \subset K$ under Assumption 6.4. From [21, Lemma A.2.2] and the Lipschitz-continuity of Π uniformly in time ensures that $\hat{\phi}$ maintain Proposition 6.8 for every pair $(x, \hat{x}) \in K \times \mathbb{R}^4$.*

Observer 6.16 operates as a hybrid system, transitioning between two dynamical systems based on the sign of $|y(t)| - \delta$. It involves states $\hat{x} \in \mathbb{R}^4$ and $\hat{v} \in \mathbb{R}^3$. The first dynamical system is a differential inclusion and is to be understood in the Fillipov sense [78]. It acts as an observer of (6.11), with output $y = x_0$ and trajectories remaining in the compact set

K (in particular $|x_0| > \delta$). This includes trajectories such that which $x(t) = G_t(v(t))$, with v solution of (6.5). Assume at some time t_1, $|y(t)| - \delta$ switches from positive to negative. Then we switch the role of $\hat{v}$ and $\hat{x}$: the trajectory of $\hat{v}$ is continued by follow the dynamics of (6.5) with initial value at t_1^+ given by $\mathcal{G}_{t_1^-}(\hat{x}(t_1^-))$. The same operation is reversed if at some time t_2, $|y(t)| - \delta$ switches from negative to positive. Under Assumption 6.4, this piecewise definition is guaranteed to be well posed (no Zeno effect). Indeed Proposition 6.13 in appendix guarantees that solutions v of (6.5) have v_0 strictly increasing near 0. In other words, there exists $\delta^* > 0$ such that trajectories remain in $\{v \in \mathbb{R}^3, |v_0| \leq \delta\}$ for a finite (and tunable) time t_δ when $\delta \leq \delta^*$. We note by $t_1 := \inf\{t \geq 0; y(t_1) = -\delta\}$ and $t_2 := \inf\{t \geq 0; y(t_2) = \delta\}$ the two possible switching time under Assumption 6.4. The value of δ^* is computed in [8, Proposition 3.4], we leave it as a black box. We give the following proposition on the existence of global solutions.

Proposition 6.10. *Suppose Assumption 6.4 holds. There exist $\delta^* > 0$ such that for any $\delta < \delta^*$, observer (6.16) admits piecewise continuous solutions (denoted $\hat{v}$ and $\hat{x}$ for simplicity) globally defined.*

Proof. The set valued function $\operatorname{sign}(\cdot)$ is upper-semi continuous, has convex and compact values. From Proposition 6.8 and Remark 6.9, as well as[78], we deduce that there exist absolutely continuous solutions to the first half of (6.16), that we denote $\hat{z}(t)$ (abusing notations for simplicity). From the absolute continuity of $\mathcal{G}$, the same holds for $\hat{v}$. The same is true for the second equation as f is globally Lipschitz 6.14. From Proposition 6.13 in appendix, there exist $\delta^* > 0$ such that if $\delta < \delta^*$ there exists at most two times $t_2 > t_1 \geq 0$ such that $y(t_1) = -\delta$ and $y(t_2) = \delta$. Hence, $\hat{v}$ and $\hat{z}$ are piece-wise continuous with, respectively, at most one jump discontinuity; at time t_1 for $\hat{z}$ and at time t_2 for $\hat{v}$. From the construction of $\mathcal{G}$, we have indeed that $\hat{v}(t_2^+) = \mathcal{G}_{t_2^-}(G_{t_2^-}(\hat{v}(t_2^-)) = v(t_2^-)$ as $|y(t_2)| \geq \delta$ which concludes the proof. ☐

We can now state the main convergence result of this paper

Theorem 6.11. *Suppose Assumption 6.4 holds. Let $0 < \delta < \delta^*$. There exist gains $k_i > 0$ and $L^* > 0$ such that for every $L \geq L^*$ there exists a class $\mathcal{KL}$ function $\mathcal{B}$ such that for every solutions of (6.16) initialised with $\hat{v}(0) \in B_{\mathbb{R}^3}(0, R)$, $\hat{z}(0) = G_0(\hat{v}(0))$ we have*

$$|v(t) - \hat{v}(t)| \leq \mathcal{B}(|v(0) - \hat{v}(0)|, t),\ t \geq t_\delta. \tag{6.17}$$

Proof. We denote by $x(t) = G_t(\bar{v}(t))$ for $t \geq 0$. Let $S := \{t \geq 0, |y(t)| \leq t_\delta\}$. From Proposition 6.13, either $S = [t_1, t_2]$, $0 \leq t_1 \leq t_2 \leq t_1 + t_\delta$, or $S = \emptyset$, in which case we say $t_1 = \infty$. From Proposition 6.10, we are assured that trajectories $\hat{v}$ and $\hat{x}$ are defined globally when $\delta < \delta^*$. We denote by $L_1 > 0$ and $L_2 > 0$ Lipschitz constants (uniform in time for L_1) of, respectively, G_t and f on $\mathbb{R}^3$.

We describe first the error during each phase of the hybrid system. When $|y| \geq \delta$ and x remains in to K, we are in position to apply [26, Proposition 4] which gives the existence of gains k_i and L^* and a class $\mathcal{KL}$ function $\mathcal{B}_1$ such that for any interval $[t_0, t_f)$ where $|y(t)| \geq \delta$ we have for all $t \in [t_0, t_f)$

$$|\hat{z}(t) - z(t)| \leq B_1(|\hat{x}(t_0) - x(t_0)|, t - t_0), \tag{6.18}$$

(we omit the fact t_0 may be a switching time, in which case t_0^+ should be employed). Using then Lemma 6.6 and Remark 6.9, we have for all $t \in [t_0, t_f)$

$$|\hat{v}(t) - v(t)| \leq \omega \circ B_1(|\hat{x}(t_0) - x(t_0)|, t - t_0). \tag{6.19}$$

If $t_0 = t_2$, i.e., we are at a switching time, then from $\hat{z}(t_2^+) = G_{t_2^-}(\hat{v}(t_2^-))$ we get for $t \in [t_0, t_f)$

$$|\hat{v}(t) - v(t)| \leq \omega \circ B_1(L_1|\hat{v}(t_0) - v(t_0)|, t - t_0). \tag{6.20}$$

Using Grönwall's lemma, the error for $t \in S$ is simply given by

$$|\hat{v}(t) - v(t)| \leq e^{L_2(t-t_1)}|\hat{v}(t_1) - v(t_1)| \leq e^{L_2 t_\delta}|\hat{v}(t_1) - v(t_1)|. \tag{6.21}$$

We now prove the existence of a global class $\mathcal{KL}$ function as stated. We consider two cases: $t_1 = 0$ and $t_1 > 0$. The case where $t_1 = \infty$ is treated similar to $t_1 = 0$. In the first case, we are assured that (6.20) holds for $t_0 = t_2$ and $t_f = \infty$ and (6.21) for $t_0 = 0$ and $t_f = t_2$. Combining the two yields for all $t \geq t_\delta$

$$|\bar{v}(t) - \hat{v}(t)| = \leq \omega \circ B_1(L_1 e^{L_2 t_\delta}|\hat{v}(0) - v(0)|, t - t_\delta), \tag{6.22}$$

which is a $\mathcal{KL}$ function as ω is a $\mathcal{K}$ function 6.6. Up to taking $L_1 \geq 1$, we denote $B_2 := \omega \circ B_1(L_1 e^{L_2 t_\delta}\cdot, \cdot)$. One can easily check that we can always extend B_2 into a new $\mathcal{KL}$ function $\bar{\mathcal{B}}$ such that it verifies the wanted point in this case.

Consider now the second case $t_1 > 0$. Combining (6.20) and (6.21) yields for all $t \geq 0$

$$|v(t) - \hat{v}(t)| \leq \omega \circ \mathcal{B}'(|v(0) - \hat{v}(0)|, t, t_2), \tag{6.23}$$

with $\mathcal{B}'(r, t, t_2)$ given by either of three cases:

$$\begin{cases} B_1(r,t) & \text{if } 0 \leq t \leq t_2 - t_\delta, \\ L_1 e^{L_2 t_\delta} B_1(r, t_2 - t_\delta) & \text{if } t_2 - t_\delta \leq t \leq t_2, \\ B_1(B_2(r, t_2 - t_\delta), t - t_2) & \text{if } t > t_2. \end{cases} \tag{6.24}$$

(We used the $\mathcal{KL}$ character of B_1, B_2 to condense dependence in t_1 and t_2 into a dependence in t_2 of the upper bound). To conclude, we prove that $\bar{\mathcal{B}}'(r,t) := \sup_{t_2 \geq 0} \mathcal{B}'(r, t, t_2)$ is a $\mathcal{KL}$ function. Continuity and growth are immediate, let us just check that $\lim_{t\to\infty} \bar{\mathcal{B}}'(r,t) = 0$. We prove it by hand. Fix $r, \varepsilon > 0$. There exists A such that we have at once

$$L_1 e^{L_2 t_\delta} B_1(r, A - t_\delta) \leq \varepsilon, \tag{6.25}$$
$$B_1(B_2(r, A - t_\delta), 0) \leq \varepsilon, \tag{6.26}$$
$$B_1(B_2(r, 0), A) \leq \varepsilon. \tag{6.27}$$

Take $t \geq 2A$. If $t_2 \geq 2A$ (6.25) implies that $\mathcal{B}'(r, t, t_2) \leq \varepsilon$, if $A \leq t_2 \leq 2A$, (6.26) implies that $\mathcal{B}'(r, t, t_2) \leq \varepsilon$, and finally if $t_2 \leq t_2$, (6.26) implies the same. Hence if $t \geq 2A$, $\bar{\mathcal{B}}'(r,t) \leq \varepsilon$.

We conclude the proof by taking $\mathcal{B} := \sup\{\bar{\mathcal{B}}', \bar{\mathcal{B}}\}$. □

6.5 INPUT-OUTPUT STABILITY ANALYSIS

As was explained in Section 6.2, Blumenfeld et al. model originates from a projection on the first few modes of an infinite dimensional neural fields model (6.2). Assuming the operator J is indeed reduced to its modes, the dynamical model for which we designed an observer, with $V^{\perp} = 0$, can be assumed to naturally occur when the input $I^{\perp}$ assumed to be null and $V^{\perp}$ converges towards its stable equilibrium at 0. Nevertheless, it is interesting to discuss the properties of our observer when the hypothesis $V^{\perp} = 0$ may not be exactly (but still closely) matched, either because $V^{\perp}(0) \neq 0$, $I^{\perp}$ is small but not null, or the operator J is not reduced to its first modes. In all generality, we consider in this section the presence of bounded disturbances on the dynamic $w = (w_0, w_1, w_2) \in L^{\infty}(\mathbb{R}_+; \mathbb{R}^3)$ and on the measurement $\nu \in L^{\infty}(\mathbb{R}_+; \mathbb{R})$ which verify respectively the Caratheodory conditions. We discuss how the noise affects observer (6.16).

In the presence of noise, system (6.5) becomes

$$\begin{cases} \dot{v}_0 = -v_0 + J_0 \Gamma_0^0(v_0, |\bar{v}|) + I_0 + w_0, \\ \dot{\bar{v}} = -\bar{v} + J_1 \Gamma_0^1(v_0, |\bar{v}|) \frac{v}{|\bar{v}|} + \bar{I} + \bar{w}, \\ y = v_0 + \nu. \end{cases} \tag{6.28}$$

The model disturbances enter linearly in the triangular system (6.11) induced by the embedding G_t,

$$\begin{aligned} \dot{x}_0 &= \phi_0(x_0, t) + x_1 + \tilde{w}_0, \\ \dot{x}_1 &= \phi_1(x_0, x_1, t) + x_2 + \tilde{w}_1, \\ \dot{x}_2 &= \phi_2(x_0, x_1, x_2, t) + x_3 + \tilde{w}_2, \\ \dot{x}_3 &= \phi_3(x_0, x_1, x_2, x_3, t) + \tilde{w}_3, \end{aligned} \tag{6.29}$$

with $\tilde{w}(t) := \partial_v G_t(v) w$, which remains uniformly bounded in time as solutions are bounded and we are under Assumption (6.4).

As the switching condition in observer (6.16) depends on the output of the system, for large disturbances on the output, Zeno phenomena may appears and the observer fails. However, it possesses some robustness to noise "small" disturbances.

Indeed, if there exist $\bar{c} > 0$ such that $I_0 - \|w_0\|_{L^\infty} > \bar{c}$ then from (6.28), using that $|\Gamma_0^0(v_0, |\bar{v}|)| \geq -\sigma(v_0)$ in the dynamics of v_0 yields that there exist $\bar{\delta}^* > 0$ such that if $|v_0| < \delta^*$, then the derivative v_0 is strictly positive. It is then clear to see that for any solution of (6.28) if $\|\nu\|_{L^\infty} < \bar{\delta}^*/2$ then if there is a time $t' \geq 0$ such that $|v_0(t) + \nu(t)| < \bar{\delta}^*/2$ then $v_0(t) + \nu(t) \geq \bar{\delta}^*/2$ for each time $t \geq t' + t_{\bar{\delta}^*/2}$, with $t_{\bar{\delta}^*/2}$ given by the expression in Proposition (6.13). We define $\bar{t}_1 := \inf\{t \geq 0, y(t) \geq -\bar{\delta}^*/2\}$ and $\bar{t}_2 := \inf\{t \geq 0, y(t) \geq \bar{\delta}^*/2\}$.

For any trajectory v of (6.28) such that $\bar{t}_1 = 0$, we follow the same idea as in the proof of Theorem 6.11. In particular, [26, Proposition 4] ensures the existence of a ISS[3] property for the x-observer in (6.16). This requires the same condition as above, mainly $I_0 - \|w_0\|_{L^\infty} > \bar{c}$ and

3 Input to state stability

$\|\nu\| < \bar{\delta}^*/2$. Then there exist $\bar{\mathcal{B}}_1$, a $\mathcal{KL}$ function, and $\gamma(\cdot,\cdot)$, a $\mathcal{K}$ function in each variable, such that

$$|\hat{x} - x| \leq \bar{\mathcal{B}}_1(|e(0)|, t) + \gamma(\|w\|_\infty, \|\nu\|_\infty). \tag{6.30}$$

This implies

$$|\hat{v} - v| \leq \omega(\bar{\mathcal{B}}_1(|e(0)|, t) + \gamma(\|w\|_\infty, \|\nu\|_\infty)), \tag{6.31}$$

with $e(0) = \hat{x}(0) - x(0)$.

The most problematic cases arise when $\bar{t}_1 > 0$. For any time $t \geq 0$ in the switch interval $[\bar{t}_1, \bar{t}_2]$, one can obtains the following estimation

$$|\hat{v}(t) - v(t)| \leq e^{L_2(t-t_1)}|\hat{v}(\bar{t}_1) - v(\bar{t}_1)| + \frac{\|w\|_\infty}{L_2}, \tag{6.32}$$

and therefore

$$|\hat{x}(t) - x(t)| \leq L_1 e^{L_2(t-t_1)}\omega(\bar{\mathcal{B}}_1(|e(0)|, t_1) + \gamma(\|w\|_\infty, \|\nu\|_\infty)) + \frac{\|w\|_\infty}{L_2}, \tag{6.33}$$

with $\bar{t}_2 - \bar{t}_1 \leq \bar{\delta}^*/2$. Let $C(t, e(0), t_1, w, \nu)$ denote the right hand side of the previous relation. We can then treat the situation after the switching time $\bar{t}_2$ by adapting (6.30). Again, from the same proposition in [26], we have for $t \geq t_2$

$$|\hat{x}(t) - x(t)| \leq \bar{\mathcal{B}}_1(|C(t_2^-, e(0), \bar{t}_1, w, \nu|, t - \bar{t}_2) + \gamma(\|w\|, \|\nu\|). \tag{6.34}$$

The estimation of $|\hat{v}(t) - v(t)|$ for $t \geq \bar{t}_2$ can then be obtain by applying ω to (6.34), and using that $|e(0)| \leq L_1|\hat{v}(0) - v(0)|$.

This description of the bound of the error is still preliminary, and does not allow to answer fully to the modelization questions we raised in the beginning of the section. Nevertheless, something can be said in the particular example where the only source of model noise comes from the fact that $V^\perp(0) \neq 0$. In that case, since $V^\perp$ converges to 0, equations (6.30)-(6.33)-(6.34) yield that convergences of the observer is still true.

6.6 NUMERICAL SIMULATIONS

We propose a numerical simulation for the observer (6.16). The parameters of the dynamical system (6.3) were chosen as such: $J_0 = -1$, $J_1 = 1.5$, $\sigma = \tanh(\mu\cdot)$ with $\mu = 10$. The distribution P has been chosen as the Dirac mass $P = \delta_{r=1}$ (corresponding to to the ring model, see [162]). The input has been chosen as $I_0(t) = \varepsilon(1-\beta)$, $\bar{I}(t) = \beta\varepsilon\big(\cos(\frac{2\pi}{10}t), \sin(\frac{2\pi}{10}t)\big)$, with $\beta = 10^{-1}$ and $\varepsilon = 10^{-1}$. The initial condition was taken as $v(0) = (-6, 2.5, -2)$ so that the output verifies $y(0) < -\delta$ to show the switching mechanism. The relaxation time constant τ has been chosen as $\tau = 5$ (for an academic case figure) to better show the error increase during the switch.

The initial condition of the observer was taken as $\hat{v}(0) = (-0.01, 2, -1)$, $\hat{x}(0) = G_0(\hat{v}(0))$. As $J_0 < 0$, we chose $\delta = 0.7$, see [8]). The gains k_i were chosen according to [26, 118]. We chose $L = 3$. The numerical scheme chosen was an explicit Euler with a step of 10^{-5}. We ran the

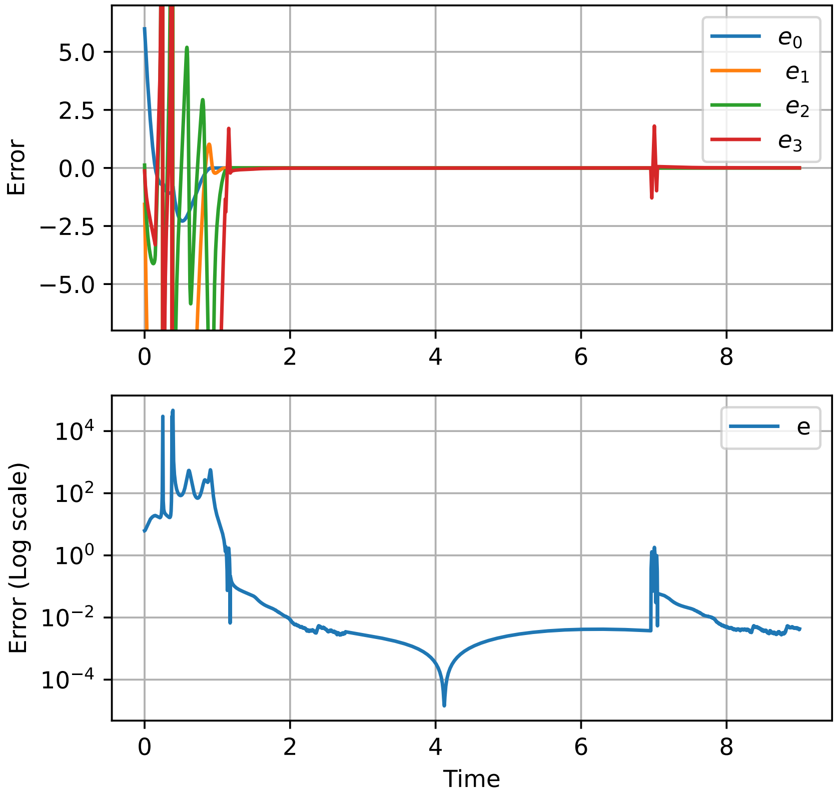


Figure 6.1: Plot of the observer error. Top: plot of $e_i(t) = \hat{x}(t) - x(t)$; bottom: plot of the error $|e|$ in log scale. The corresponding trajectories v was taken as such $v(0) < \delta$. The switching time occur at $t_1 \simeq 2.5$ and $t_2 \simeq 6.5$. Peaking phenomena occur at the start of the simulation and at time t_2, when the sliding mode observer in x restarts; they are better seen on the log-error plot. Since the error peaking at t_2 is proportional to the error at t_1, it is much smaller in size which show the point in introducing the hybrid system. At time t_1, the sliding mode observer is turned off and the corresponding trajectories of v and $\hat{v}$ evolve according to the neural fields dynamics. x and $\hat{x}$ are tracked through the mapping G, resulting in a drift of the error over $[t_1, t_2]$.

simulation in the interval $[0, 9]$. We show the trajectories on the immersed domain for better clarity. The errors $e_i(t) := \hat{x}(t) - x(t)$ and the error $|e| = |\hat{x}(t) - x(t)|$ in log scale are shown in Figure 6.1. We added another Figure 6.2 to show the dynamic of the error before the switch and highlight the sliding mode.

APPENDIX: COMPUTATIONAL LEMMAS

Computational lemmas

We recall the following results proved in [8].

Lemma 6.12. *The following assertions hold.*

1. *The quantities Γ_p^j are well defined for all $j, p \in \mathbb{N}$ and belong to the class $C^\infty(\mathbb{R} \times [0, \infty), \mathbb{R})$.*
2. *We have the following relations*

$$\partial_{v_0} \Gamma_p^j = \Gamma_{p+1}^j \quad \text{and} \quad \partial_\rho \Gamma_p^j = \Gamma_{p+1}^{j+1}, \qquad \forall j, p \in \mathbb{N}. \tag{6.35}$$

3. *For any $v_0 \neq 0$, the function $\Gamma_0^0(v_0, \cdot) \in C^\infty([0, \infty), \mathbb{R})$ is injective.*
4. *For any $v_0 \neq 0$ and $\rho > 0$, we have $\Gamma_1^1 \neq 0$ and of opposite sign w.r.t. v_0. Moreover, $\Gamma_1^1(v_0, 0) = 0$ for any $v_0 \in \mathbb{R}$.*
5. *For any $\rho \geq 0$, we have $\Gamma_0^0(0, \rho) = 0$.*

Proposition 6.13. *Consider an input $I : [0, +\infty) \to \mathbb{R}^3$ such that $I_0(t) \geq c > 0$ for all $t \geq 0$, and a solution $\bar{v}$ of (6.3) associated to this input. Let*

$$\delta^* = \frac{c}{1 + |J_0| \sigma'(0)}, \tag{6.36}$$

and fix some δ, $0 < \delta < \delta^*$. The first component $v_0(t)$ of the solution vanishes at most one time, and the set of times t such that $|v_0(t)| \leq \delta$ is a time-interval of length no larger than t_δ given by

$$t_\delta = \frac{\delta^*}{c} \frac{2\delta}{\delta^* - \delta}. \tag{6.37}$$

Moreover, if there exist a time $t^* \geq 0$ such that $v_0(t^*) \geq \delta$, then $v(t) \geq \delta$ for all times $t \geq t^*$.

Regarding the existence of solutions of (6.3) and an invariant set, we have the following from [8]

Proposition 6.14. *Assume $I \in C^0([0, +\infty))$. The dynamics (6.3) admit a unique global solution v defined on $\mathbb{R}$ for every initial condition $v(0) \in \mathbb{R}^3$. Moreover, if $\sup_t \|I\| < \infty$, then there exist $R^* > 0$, such that for any $R \geq R^*$, $B_{\mathbb{R}^3}(0, R)$ is an invariant attracting set by the dynamic f.*

Table 6.1: Hölder restrictions on ϕ for a sliding mode observer (on compact sets). See [21, 26].

	$i \backslash j$	1	2	3	4
	1	$\frac{3}{4}$			
$\alpha_{ij} \geq$	2	$\frac{1}{2}$	$\frac{2}{3}$		
	3	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	
	4	0	0	0	0

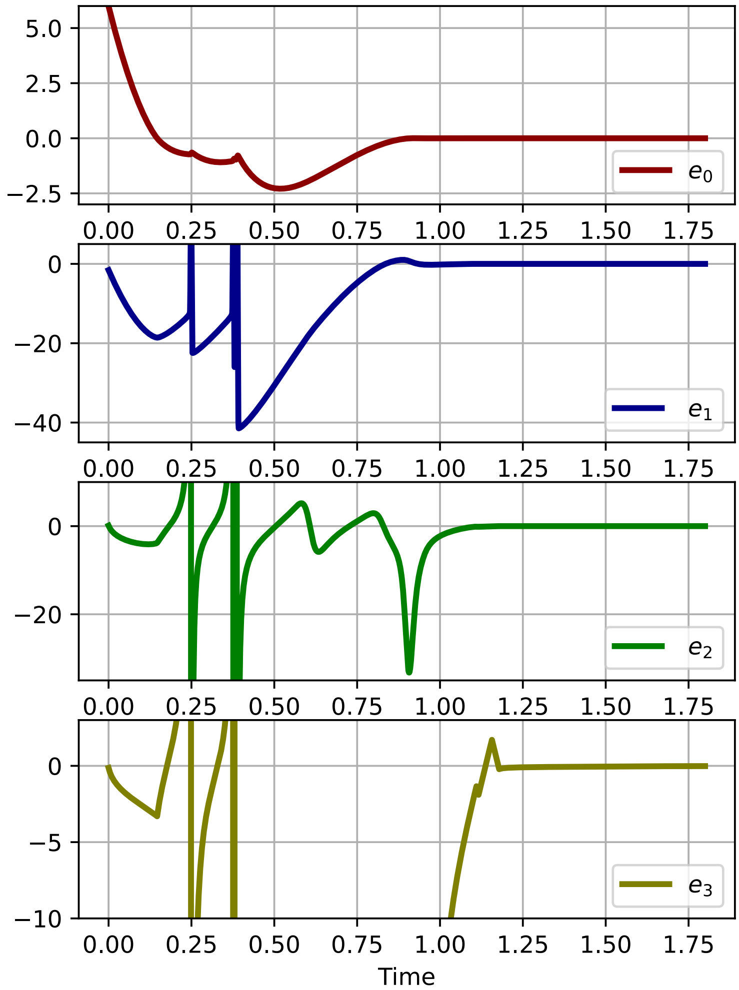


Figure 6.2: Components of the error $e_i(t) = x(t) - \hat{x}(t)$ before the switching time t_1. We see the sliding effect on the trajectory e_4 in the bottom graph and the overall finite time convergence.

Part III

PARAMETER ESTIMATION OF SATURATED SYSTEMS

This part addresses the problem of parameter estimation for networks of neural mass models with saturated interconnections. The central class of systems considered is the two-population network

$$\begin{cases} \dot{V}_1 = A_1 V_1 + W_{11} S(V_1) + W_{12} S(V_2) + u_1 + h_1, \\ \dot{V}_2 = A_2 V_2 + W_{21} S(V_1) + W_{22} S(V_2) + u_2 + h_2, \\ y = V_1, \end{cases}$$

where only the first population V_1 is measured. The saturation function S and the inputs u_i are known. All connection matrices W_{11}, W_{12}, W_{21}, W_{22}, the internal dynamics A_1, the bias terms h_1, h_2, and the hidden state V_2 are unknown. Here, the nonlinear function S is a function that exhibits saturation, such as a sigmoid or a piecewise linear function.

The goal is to design input signals u_1, u_2 and algebraic estimators that recover these parameters exactly and in finite time from the input-output data (u, y).

The next chapter introduces the main difficulties in a simplified scalar case and develops the core estimation tools that underpin the full protocol. The subsequent chapters develop then the tools in both finite and infinite-dimensional settings, with increasing complexity of the treated systems.

In Chapter 11 we present the full protocol for the two-population network which achieves finite time estimation of the desired parameters and the state V_2.

Lastly, we extend this result to hierarchical networks of N populations where only the first one is measured:

$$\begin{cases} \tau_1 \dot{V}_1 = -V_1 + W_{1,1} S(V_1) + W_{1,2} S(V_2) + u_1, \\ \tau_2 \dot{V}_2 = -V_2 + W_{2,1} S(V_1) + W_{2,2} S(V_2) + W_{2,3} S(V_3) + u_2, \\ \quad \vdots \\ \tau_{N-1} \dot{V}_{N-1} = -V_{N-1} + W_{N-1,1} S(V_1) + \cdots + W_{N-1,N} S(V_N) + u_{N-1}, \\ \tau_N \dot{V}_N = -V_N + W_{N,1} S(V_1) + \cdots + W_{N,N-1} S(V_{N-1}) + W_{N,N} S(V_N) + u_N, \\ y = V_1. \end{cases}$$

We show in chapter 11 that all matrices W_{ij} and hidden states V_i of this system can be identified in finite time with a specific protocol of input signals and estimators, which is an extension to the two-layer case.

7

TWO DIMENSIONAL SYSTEMS

This chapter starts the third part of the thesis, devoted to building algorithms for joint state and parameter estimation of neural mass systems or networks of such systems; the complexity of the treated systems increases with chapters. The present chapter, being the first chapter of Part III, considers the simplest case possible (two layers of dimension 1), and is devoted on the one hand to explaining the difficulty of our objective, due to terms where unknown parameters are multiplied by a function of the unknown state, and on the other hand to demonstrating, in this simpler framework, the basic building blocks of the solution we are proposing throughout Part III.

7.1 THE CLASS OF SYSTEMS AND THE ESTIMATION PROBLEM

The whole chapter revolves around the two-dimensional system

$$\begin{cases} \dot{x}_1 = \phi_1(x_1) + w_2\, S(x_2) + w_0, \\ \dot{x}_2 = \phi_2(x_1, x_2, t) + u, \\ y = x_1, \end{cases} \tag{7.1}$$

where $x = (x_1, x_2) \in \mathbb{R}^2$, $u \in \mathbb{R}$ is a control input and $y = x_1$ is the only measurement. The function $\phi_1 : \mathbb{R} \to \mathbb{R}$ is known and globally Lipschitz. The function $\phi_2 : \mathbb{R}^2 \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ is unknown and only assumed to be continuous; the parameters $w_0, w_2 \in \mathbb{R}$ are unknown scalars with a known a priori bound $|w_0|, |w_2| \leq W$. The map $S : \mathbb{R} \to [0, 1]$ is usually supposed to be a sigmoid function, specifically, a smooth bounded strictly increasing function, one example being the logistic function $S(x) = 1/(1 + e^{-\lambda x})$ with slope $\lambda > 0$.

The objective is to reconstruct $(w_0, w_2, x_2(t))$ jointly, in finite time, from the past input $u_{[0,t]}$ and the past output $y_{[0,t]}$. System (7.1) is the elementary case of a measured population driven by an unmeasured one through a sigmoidal coupling weighted by an unknown synaptic gain, the very structure that underlies neural mass models based on Wilson-Cowan equations [168, 167] and reappears, recursively, in the cascades and connectivity matrices of large-scale brain models [36, 58, 142, 97, 102].

7.2 WHY CLASSICAL METHODS DO NOT WORK HERE

Before introducing the new method, it is useful to recall the standard paradigm in the parameter-estimation literature and to see precisely where it breaks down for (7.1). In its simplest form, parameter estimation starts from a regression in which the unknown parameter enters linearly and the regressor is available from measurements or from known signals. A typical model is [121, 147, 172]

$$y(t) = \varphi(t)^\top \theta + \varepsilon(t), \tag{7.2}$$

where y and φ are known from data and θ is unknown. The least-squares estimate over $[0, t]$ is

$$\hat{\theta}_{\mathrm{LS}}(t) = M(t)^{-1} \int_0^t \varphi(\tau) y(\tau)\, d\tau, \qquad M(t) := \int_0^t \varphi(\tau)\varphi(\tau)^\top d\tau,$$

whenever $M(t)$ is invertible, and gradient or recursive variants are obtained by differentiating the same formula online:

$$\dot{\hat{\theta}} = P(t)\, \varphi(t)\big(y(t) - \varphi(t)^\top \hat{\theta}\big), \qquad \dot{P} = -P\, \varphi(t)\varphi(t)^\top P, \quad P(0) = P_0 > 0. \tag{7.3}$$

Modern variants such as DREM [11, 132] or concurrent learning [49, 50] relax the excitation requirement, but none removes the prerequisite, the regression must be available and the regressor $\varphi(t)$ must be computable.

When $\varphi(t)$ depends on a state that is not fully measured, the regression cannot be formed directly. *Most* adaptive observers are introduced precisely to close this gap [29]. A first systematic way to do this, due to Zhang [171] and Besançon [31], is to seek a change of variables that puts the system into form

$$\begin{cases} \dot{z} = Z_0(z, y, u, t), \\ y = H(z, u, t) + G(z, u, t)\, \theta, \end{cases} \tag{7.4}$$

where all functions are known and the hidden dynamics for z do not depend on θ. The key observation is that if z were measured, the output equation would already be a linear regression in θ. Therefore the two main assumption is that we reconstruct z which is granted by assuming that the dynamics on z are contracting, allowing asymptotic reconstruction. The second is to ensures that alongside its trajectories $G(z, u, t)$ is computable and verify some excitation condition.

A second route, applicable to triangular systems, does not require the parameter to appear in the output at all. It is enough that it enter the last state equation of a triangular chain where the parameter appears either sequentially or only at the last equation:

$$\begin{cases} \dot{\xi}_1 = \xi_2 + \phi_1(\xi_1, u), \\ \vdots \\ \dot{\xi}_n = \phi_n(\xi, u) + \psi(\xi, u)\, \theta, \\ y = \xi_1, \end{cases} \tag{7.5}$$

with all ϕ_i, ψ known. This can be extended where there are n parameters θ_i appearing successively at each layers as authors have shown in [77]. Under appropriate observability and excitation conditions, state and parameter estimation can be carried out simultaneously [30]. The intuition is clearest through the approach of [80], where a cascade of sliding-mode observers is used to reconstruct the state ξ and the term $\psi(\xi, u)\,\theta$ from the output in finite time. Then assuming that $\psi(\xi, u)$ is computable and verify some excitation condition, the parameter θ can then be estimated.

For system (7.1), the obstruction appears already at the level of the measured equation. Writing

$$Z(t) := w_2 S(x_2(t)) + w_0, \tag{7.6}$$

the first state equation becomes $\dot{x}_1 = \phi_1(x_1) + Z(t)$. Since x_1 is measured and ϕ_1 is known, one may construct an exact differentiator to reconstruct $Z(t)$ for all t. Then, the identification problem reduces to

$$Z(t) = w_2\, v(t) + w_0, \qquad v(t) = S(x_2(t)), \tag{7.7}$$

in which the natural regressor is $v(t) = S(x_2(t))$. This regressor is not measured, and the decomposition itself is not unique. Indeed, for any scalar $\lambda > 0$, the substitution $(w_2, v) \mapsto (\lambda w_2, v/\lambda)$ leaves Z unchanged. So the information contained in Z alone is insufficient to separate the parameter w_2 from $S(x_2(t))$. Furthermore, as the dynamic in x_2 is unknown, classical approaches of adaptive observers explained above cannot be used to reconstruct x_2. If for example the dynamic on x_2 were contracting, one could reconstruct x_2 asymptotically and we would be in the case of (7.4) [171]. Recent adaptive observers proposed for neural-mass and neural-field models [40, 41, 44] use this contraction assumption to solve this ambiguity.

As the problem is not well-posed in the classical framework for any injective function S, we *slightly* relax the assumption on S. Indeed, as it is in most models a bounded and strictly increasing function, it gets closer and closer to its bound, to then be *almost* equal to it above a certain threshold. A common approach in the mathematical neuroscience community is to assume that there exists an upper and lower threshold, $(\bar{R}, \underline{R}) \in \mathbb{R}^2$ respectively such that $S(x) = 1$ or $S(x) = 0$ for $x \geq \bar{R}$ or $x \leq \underline{R}$ (see for example [54]). One can then quantify the small error between the classical sigmoid and this idealized saturation function, which becomes a bounded perturbation of the problem that can be made as small as desired by choosing the thresholds appropriately.

The question now is therefore: Can we exploit the very structure of the nonlinearity S to lift this ambiguity on Z, to render $S(x_2)$ effectively measurable without even knowing x_2 nor its dynamics ϕ_2, barring a known bound, and at the same time render it sufficiently rich to allow the estimation of w_2 by building on the classical framework? Can we construct a *procedure*, namely a control u and an estimation algorithm, that achieves this goal? The next section answers these questions in the affirmative, by showing how to exploit the saturation structure of S to separate w_2 from $S(x_2)$ and to reconstruct x_2 without knowing ϕ_2 beyond a known bound.

7.3 NOVEL ESTIMATION METHOD VIA SATURATION FUNCTIONS

We recall the system under study:

$$\begin{cases} \dot{x}_1 = \phi_1(x_1) + w_2\, S(x_2) + w_0, \\ \dot{x}_2 = \phi_2(x_1, x_2, t) + u, \\ y = x_1, \end{cases} \tag{$\star$}$$

where $w_0, w_2 \in \mathbb{R}$ are unknown parameters, and ϕ_2 is an unknown continuous nonlinearity. The goal is the joint estimation of w_0, w_2, x_2 despite the unknown ϕ_2. The function $S : \mathbb{R} \to \mathbb{R}$ is a known saturation-type nonlinearity satisfying the following key assumption.

Remark 7.1. *The term w_0 can be replaced by a time-varying function $g(t)^\top\theta$, where g is a known persistently exciting function and $\theta \in \mathbb{R}^m$ is an unknown parameter vector. Then the objective becomes estimating θ. All the results discussed below still hold, because in the estimation process we typically obtain an estimate $\hat{w}_0$ over a long interval $\mathcal{I}$ (i.e., $\hat{w}_0 = w_0\mathbf{1}_{\mathcal{I}}$). If we replace $w_0(t)$ by $g(t)^\top\theta$, one can then deduce θ from the estimate of w_0. The choice of a constant w_0 is made to simplify the proof in this introductory exposition. The main theorem with the full generalization will be given in the next chapter.*

Assumption 7.2 (Saturation Function)**.** *The function $S : \mathbb{R} \to \mathbb{R}$ is a known, globally Lipschitz-continuous, piecewise continuously differentiable function. It is characterized by known saturation levels $\underline{R} < \bar{R}$ such that:*

$$x \leq \underline{R} \Rightarrow S(x) = 0 \quad \textit{and} \quad x \geq \bar{R} \Rightarrow S(x) = 1, \tag{7.8}$$

and is strictly increasing on the interval $[\underline{R}, \bar{R}]$. Consequently, $\|S\|_\infty = 1$. See Figure 7.1. For vector arguments $x \in \mathbb{R}^n$, the function is applied component-wise:

$$S(x) = (S(x_1), \ldots, S(x_n))^\top$$

Assumption 7.3 (Admissible hidden dynamics)**.** *The function $\phi_2 : \mathbb{R}^2 \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ is continuous, and one of the following two alternatives holds:*

1. *There exists a known constant $m > 0$ such that $\|\phi_2\|_\infty \leq m$.*
2. *There exist known constants $\lambda, m > 0$ and a continuous function ψ with $\|\psi\|_\infty \leq m$ such that*

$$\phi_2(x_1, x_2, t) = -\lambda x_2 + \psi(x_1, x_2, t).$$

More generally, it is enough that for each bounded control used below, the resulting trajectories remain in a known invariant attracting compact set on which ϕ_2 is bounded by a known constant.

Remark 7.4. *The saturation function defined in Assumption 7.2 represents an idealization that differs from the common smooth sigmoidal functions, which are strictly increasing everywhere. This idealization yields sharper analytical results and is standard in the literature [52]. Once these results are established in the idealized setting, we will attempt to extend them to the classical case of smooth sigmoidal functions by interpreting any smooth sigmoid as an idealized saturation function perturbed by a small, bounded additive term. This will be done by proving that our estimation method is robust to bounded noise.*

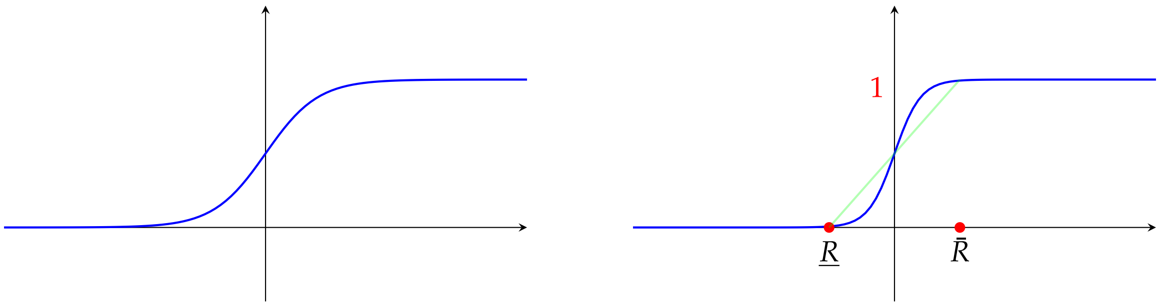


Figure 7.1: Comparison of two types of activation functions. On the left side : Typical smooth sigmoidal function. On the right side : Idealized saturation with explicit thresholds.

The structure of the non-linear function S opens a novel path for estimation. To illustrate the core idea, consider a simplified static estimation problem. Suppose we have as measure a signal $\tilde{Y}(t)$ of the form:

$$\tilde{Y}(t) = \tilde{w}\, S(\tilde{v}(t)) + \tilde{w}_0, \tag{7.9}$$

where $\tilde{w}$, $\tilde{w}_0$, and the signal $\tilde{v}(t)$ are all unknown and we want to estimate $\tilde{w}$ and $\tilde{w}_0$.

In the linear case, where the measured signal is of the form $Y(t) = wX(t) + w_0$ with some signal unknown X, estimating the parameters w and w_0 requires, at a minimum, the existence of two distinct time instants $t_1, t_2 > 0$ at which the input signal $X(t)$ takes known and distinct values $X(t_1) \neq X(t_2)$. In the presence of noise, one typically requires full measurement of the signal $X(t)$ over a certain interval $[0, T]$ to cast the estimation as a minimization problem (e.g., least squares) over that interval. In this case, along with a persistence of excitation condition, the problem admits a unique solution w, w_0.

In contrast, in the saturated form, due to the very nature of the nonlinearity S, we can solve the estimation problem while having large uncertainties on the signal $\tilde{v}$. Indeed, if $\hat{v}$ is in the:

1. Lower saturation region i.e. If there exists a time interval $\mathcal{I}_1$ where $\tilde{v}(t) \leq \underline{R}$, then $S(\tilde{v}(t)) = 0$ on $\mathcal{I}_1$, yielding
$$\tilde{Y}(t) = \tilde{w}_0, \quad \forall t \in \mathcal{I}_1. \tag{7.10}$$

2. Upper saturation region i.e. If there exists an interval $\mathcal{I}_2$ where $\tilde{v}(t) \geq \bar{R}$, then $S(\tilde{v}(t)) = 1$ on $\mathcal{I}_2$, yielding
$$\tilde{Y}(t) = \tilde{w} + \tilde{w}_0, \quad \forall t \in \mathcal{I}_2. \tag{7.11}$$

Under the assumption that such informative intervals exist, a simple averaging estimator $E : C(\mathbb{R}, \mathbb{R}) \to \mathbb{R}^2$ can be constructed:

$$E(\tilde{Y}) = \left(\bar{Y}_{\mathcal{I}_1},\ \bar{Y}_{\mathcal{I}_2} - \bar{Y}_{\mathcal{I}_1}\right), \tag{7.12}$$

where $\bar{Y}_{\mathcal{I}} = \frac{1}{|\mathcal{I}|}\int_{\mathcal{I}} \tilde{Y}(t)\, dt$ denotes the average value over interval $\mathcal{I}$. This elementary estimator recovers indeed the parameters: $E(\tilde{Y}) = (\tilde{w}_0, \tilde{w})$.

Remark 7.5. *It is not necessarily required that the saturation S, under Assumption 7.2, takes as upper value 1 and a as lower values 0. This choice is made to simplify the computations. As long as the two* *values are known and different we can still solve the estimation problem (7.9). Indeed, suppose we had instead*

$$S(x) = s_1,\ x \geq \bar{R}, \quad S(x) = s_2,\ x \leq \underline{R}, \quad s_2 > s_1 \tag{7.13}$$

and suppose that $s_2 \neq 0$. By Assumption 7.2, at most one of the saturation levels s_1 and s_2 may be zero. Consequently, if $\tilde{v}$ is in the:

1. *Lower saturation region: If there exists a time interval $\mathcal{I}_1$ where $\tilde{v}(t) \leq \underline{R}$, then $S(\tilde{v}(t)) = s_2$ on $\mathcal{I}_1$, yielding*
$$\tilde{Y}(t) = \tilde{w}s_2 + \tilde{w}_0, \quad \forall t \in \mathcal{I}_1. \tag{7.14}$$
2. *Upper saturation region: If there exists an interval $\mathcal{I}_2$ where $\tilde{v}(t) \geq \bar{R}$, then $S(\tilde{v}(t)) = s_1$ on $\mathcal{I}_2$, yielding*
$$\tilde{Y}(t) = \tilde{w}s_1 + \tilde{w}_0, \quad \forall t \in \mathcal{I}_2. \tag{7.15}$$

The estimator in this case would be

$$E_2(\tilde{Y}) = \left(\frac{\bar{Y}_{\mathcal{I}_1} - \bar{Y}_{\mathcal{I}_2}}{s_2 - s_1}, \frac{s_2\bar{Y}_{\mathcal{I}_2} - s_1\bar{Y}_{\mathcal{I}_1}}{s_2 - s_1} \right) \tag{7.16}$$

which would recover $E_2(\tilde{Y}) = (\tilde{w}, \tilde{w}_0)$

7.3.1 *Observer design for the saturated system*

We now turn to the application of the estimation strategy outlined above (7.9) to the dynamical system ($\star$). Besides Assumption 7.3, we only require a priori bounds on the unknown parameters and on the initial condition.

Assumption 7.6 (A priori bounds)**.** *There exists a known positive constant B such that:*

- *The unknown parameters are bounded: $|w_0| \leq B$ and $|w_2| \leq B$;*
- *The initial condition is bounded: $\|x(0)\| \leq B$.*

A key step in our approach involves reconstructing the aggregate signal $w_2S(x_2) + w_0$ from the output measurements. This is achieved via a two-dimensional homogeneous sliding-mode observer that we introduced earlier. We assume the control input u is known, bounded, and piecewise continuous.

To facilitate observer design, for any solution, consider the following rewriting

$$X_2(t) := w_2S(x_2(t)) + w_0. \tag{7.17}$$

The original system ($\star$) can then be reformulated a.e as:

$$\begin{cases} \dot{x}_1 = \phi_1(x_1) + X_2(t), \\ \dot{X}_2(t) = w_2S'(x_2(t))\big(\phi_2(x_1(t), x_2(t), t) + u(t)\big) =: \Phi_2(x_1(t), x_2(t), t), \\ y = x_1, \end{cases} \tag{7.18}$$

where S' denotes the derivative of S (which exists almost everywhere due to Assumption 7.2). Crucially, under Assumptions 7.6 and 7.3, and for bounded controls u, the term Φ_2 is itself bounded by a known constant along the considered trajectories.

Leveraging the finite-time convergence properties of the super-twisting observer , presented in the first section, for systems of the form (7.18), we propose the following observer:

$$\begin{cases} \dot{\hat{z}}_1 = \phi_1(y) + \hat{z}_2 + L\left\lceil y - \hat{z}_1 \right\rfloor^{1/2}, \\ \dot{\hat{z}}_2 \in \mu L^2 \operatorname{sign}(y - \hat{z}_1), \\ \hat{Y}(t) = \hat{z}_2(t) \cdot \mathbf{1}_{\{t \geq T\}}, \end{cases} \tag{7.19}$$

where $L > 0$ is the observer gain, $\mu > 0$ is a fixed design parameter, and $T > 0$ denotes a prescribed convergence time. The indicator function $\mathbf{1}_{\{\cdot\}}$ ensures that the estimate $\hat{Y}$ is only considered valid after time T.

The convergence of this observer is formalized in the following proposition.

Proposition 7.7. *Consider system (7.18) under Assumptions 7.6 and 7.3. For any prescribed convergence time $T > 0$, for any $\rho > 0$ there exists a gain $L^* > 0$ such that for all $L > L^*$, the observer (7.19) initialized in $B_2(0, \rho)$ guarantees:*

$$\hat{Y}(t) = w_2 S(x_2(t)) + w_0, \quad \forall t \geq T. \tag{7.20}$$

Proof. The result follows directly from the finite-time stability properties of the super-twisting observer [117, 129, 26]. The boundedness of Φ_2 ensures the existence of a sufficiently large gain L that overcomes the perturbation, guaranteeing exact convergence of $(\hat{z}_1, \hat{z}_2)$ to (x_1, X_2) within time T. □

Thus, after a finite transient period T, the observer (7.19) provides exact reconstruction of the aggregate signal $X_2(t) = w_2 S(x_2(t)) + w_0$.

The subsequent challenge is to disentangle the individual parameters w_0 and w_2 from this aggregate measurement. As established, the proposed estimation strategy for the problem requires the a priori knowledge of two time intervals, $\mathcal{I}_1$ and $\mathcal{I}_2$, in which the saturation function $S(\tilde{v}(t))$ attains its upper and lower values. To achieve this, we are going to leverage the control u appearing on the dynamic of x_2 in ($\star$) together with the standing Assumption 7.3. This will be the objective of the next section

7.3.2 *Open-loop steering of the non-linearity S*

The objective of this section is to demonstrate some form of open-loop controllability of the nonlinearity $S(x_2)$. By choosing specific constant controls u we can steer the values of $S(x_2(t))$ into values 1 or 0.

We begin with the first alternative of Assumption 7.3, for which the dynamics of x_2 reduce to systems of the type

$$\dot{x} = f_0(x, t) + u, \tag{7.21}$$

with f_0 unknown continuous bounded function with known bound. Note that we can also recover the above form by recalling that in ($\star$) we measure x_1 and can simply do an output injection in ϕ_2.

The following proposition shows that assuming a boundedness condition on f_0 we can control $S(x)$ into a certain accessible set without knowledge of x

Lemma 7.8. *Let $T, \rho, m > 0$ and $h \in \mathcal{A} := \{0, 1\}$. Consider the constant control*

$$u_\mu^h := \begin{cases} \mu, & \text{if } h = 1, \\ -\mu, & \text{if } h = 0, \end{cases} \tag{7.22}$$

with $\mu \geq \mu^*$, where

$$\mu^* := m + \frac{\rho + \max\{|\bar{R}|, |\underline{R}|\}}{T}. \tag{7.23}$$

Then for any continuous f_0 such that $\|f_0\|_\infty \leq m$ solutions of

$$\dot{x} = f_0(x, t) + u^h, \tag{7.24}$$

initialised in $B(0, \rho)$ verify

$$-\rho + (\mu - m)t \leq \operatorname{sign}(u^h)x(t) \leq \rho + (m + \mu)t. \tag{7.25}$$

In particular, we have that

$$S(x(t)) = h, \quad t \geq T. \tag{7.26}$$

With, $\bar{R}$, $\underline{R}$ set in Assumption 7.2.

Proof. Let us show that if $h = 1$ then we have indeed the claim. From the bound on f_0 and using (7.22) we have

$$-m + \mu \leq \dot{x} \leq m + \mu. \tag{7.27}$$

From Grönwall's lemma and using that solutions are initialized in the ball $B(0, \rho)$ we obtain indeed

$$-\rho + (\mu - m)t \leq \operatorname{sign}(u^h)x(t) \leq \rho + (m + \mu)t. \tag{7.28}$$

Finally we can check that at time T we have

$$x(t) \geq T(\mu - m) - \rho \geq T(\mu^* - m) - \rho \geq \bar{R} \tag{7.29}$$

which indeed imply from Assumption 7.2 that $S(x(t)) = h$ for $t \geq T$. A similar argument yields the result for $h = 0$. □

From the above lemma, we immediately obtain the following

Corollary 7.3.1. *Let $T_4 > T_3 > T_2 > T_1 > 0$ and $\rho, m > 0$. Consider $h_1, h_2 \in \{0, 1\}$ with $h_1 \neq h_2$. Then there exists μ_1^* such that for any $\mu_1 > \mu_1^*$ there exists $\mu_2^* > \mu_1$ such that for every $\mu_2 > 0$ verifying $\mu_2 > \mu_2^* > \mu_1 > \mu_1^*$ and for any continuous f_0 such that $\|f_0\|_\infty \leq m$ solutions of*

$$\dot{x} = f_0(x, t) + u_{\mu_1}^{h_1}\mathbf{1}_{[0,T_2]} + u_{\mu_2}^{h_2}\mathbf{1}_{(T_2,T_4]}, \tag{7.30}$$

initialised in $B(0, \rho)$ verify

$$S(x(t)) = \begin{cases} h_1 & t \in [T_1, T_2], \\ h_2 & t \in [T_3, T_4]. \end{cases} \tag{7.31}$$

Proof. First, by Carathéodory's existence theorem, solutions of (7.30) are absolutely continuous and globally defined in time. This follows from the continuity of the right-hand side (except at the single switching instant $t = T_2$) and its boundedness on any finite interval.

Second, choosing μ_1^* according to Lemma 7.8 with parameters $T = T_1$, m, and ρ, we have that for any $\mu_1 > \mu_1^*$, the control $u_{\mu_1}^{h_1}$ ensures that $S(x(t)) = h_1$ for all $t \in [T_1, T_2]$.

At time T_2, the lemma provides an explicit bound on the solution:

$$|x(T_2)| \leq \max\{|-\rho + \mu_1 T_2|,\ |\rho + (m + \mu_1)T_2|\} =: \rho_2. \tag{7.32}$$

We now reapply Lemma 7.8 on the interval $[T_2, T_4]$ with initial radius ρ_2, time horizon $T_3 - T_2$, and the same bound m. This yields the threshold

$$\mu_2^* := m + \frac{\rho_2 + \max\{|\bar{R}|, |\underline{R}|\}}{T_3 - T_2}. \tag{7.33}$$

For any $\mu_2 > \mu_2^*$, the control $u_{\mu_2}^{h_2}$ guarantees that $S(x(t)) = h_2$ for all $t \in [T_3, T_4]$. □

We can still find a condition of f_0 that ensures open-loop controllability of the nonlinearities while ensuring that the state remains in a known compact, for ideally all times. Indeed, if f_0 admits the following structure

$$f_0(x,t) = f_1(x) + \psi(x,t), \tag{7.34}$$

with ψ continuous bounded function supposed unknown with known bound $m > 0$ and f_1 continuous unknown function verifying a passivity-type condition

$$x f_1(x) \leq -\lambda |x|^2 + \beta \tag{7.35}$$

with $\lambda > 0$ and $\beta \in \mathbb{R}$ known values, it is possible to derive results similar to Lemma 7.8 and Corollary 7.3.1 for systems with f_0 as in (7.34). This is due to the fact that solutions x of (7.34) are absolutely continuous and global in time for any initial condition in $\mathbb{R}$ and satisfy

$$\frac{1}{2}\frac{d}{dt}|x|^2 \leq -\lambda |x|^2 + (\beta + m)|x|. \tag{7.36}$$

This implies that the ball $B(0, \frac{\beta+m}{\lambda})$ is an invariant attractive set for (7.21)

A particularly relevant case is systems where f_0 takes this simplified form

$$f_0(x,t) = -\lambda x + \psi(x,t), \tag{7.37}$$

which is exactly the form that takes the Wilson-Cowan equations [167] which will be relevant when we study x in higher dimension.

For the above simplified form we have the following lemma

Lemma 7.9. *Let $T > 0$, $\rho, m, \lambda > 0$, and $h \in \mathcal{A} := \{0, 1\}$. Consider the constant control u_μ^h*

$$u_\mu^h := \begin{cases} \mu, & \text{if } h = 1, \\ -\mu, & \text{if } h = 0, \end{cases} \tag{7.38}$$

with $\mu \geq \mu^*$, where

$$\mu^* := m + \lambda \frac{\rho + e^{\lambda T} \max\{|\underline{R}|, |\bar{R}|\}}{e^{\lambda T} - 1}. \tag{7.39}$$

Then, for any continuous ψ such that $\|\psi\|_\infty \leq m$, solutions of

$$\dot{x} = -\lambda x + \psi(x,t) + u^h. \tag{7.40}$$

initialized in $B_n(0,\rho)$, satisfy

$$|x(t)| \leq (m+\mu)/\lambda \tag{7.41}$$

$S(x(t)) = h$ when $t \geq T$. Here, $\bar{R}$ and $\underline{R}$ are set in (7.8).

Proof. Let us show that if $h = 1$, then $x(t) \geq \bar{R}$ for all times $t \geq T$. From the expression of μ^* in (11.12), we have $\dot{x} \geq -\lambda x + (\rho + \bar{R})/(e^T - 1)$. The claim follows by Grönwall's lemma and the fact that $x(0) \geq -\rho$. A similar argument shows that $x(t) \leq \underline{R}$ if $h = 0$, completing the proof. □

The advantage of this form is that the known compact where the state lives does not depend on time, which is encapsulated in (7.25). Just as previously, we can extend the result of this form to the following corollary

Corollary 7.3.2. *Let $T_4 > T_3 > T_2 > T_1 > 0$ and $\rho, m > 0$. Consider a sequence $h_1, h_2 \in \{0,1\}$ with $h_1 \neq h_2$. Then there exists μ_1^* such that for any $\mu_1 > \mu_1^*$ there exists $\mu_2^* > \mu_1$ such that for every $\mu_2 > 0$ verifying $\mu_2 > \mu_2^* > \mu_1 > \mu_1^*$ for any continuous ψ such that $\|\psi\|_\infty \leq m$ solutions of*

$$\dot{x} = -\lambda x + \psi(x,t) + u_{\mu_1}^{h_1} \mathbf{1}_{[0,T_2]} + u_{\mu_2}^{h_2} \mathbf{1}_{(T_2,T_4])}, \tag{7.42}$$

initialized in $B(0,\rho)$ verify

$$S(x(t)) = \begin{cases} h_1 & t \in [T_1, T_2], \\ h_2 & t \in [T_3, T_4]. \end{cases} \tag{7.43}$$

The two open-loop controllability corollaries above correspond exactly to the two alternatives in Assumption 7.3: Corollary 7.3.1 applies in the purely bounded case, whereas Corollary 7.3.2 applies in the dissipative case $\phi_2(x_1, x_2, t) = -\lambda x_2 + \psi(x_1, x_2, t)$.

7.3.3 *First adaptive observer*

From the results obtained in the above section on the open loop controllability, we improve on the adaptive observer introduced in (7.19) to recovers the quantities $S(x_2), w_0, w_2$ of system ($\star$)

Let $T_4, T_3, T_2, T_1, L, \epsilon > 0$. Consider the following adaptive observer

$$\begin{cases} \dot{\hat{z}}_1 = \hat{z}_2 + L\lceil y - \hat{z}_1 \rfloor^{\frac{1}{2}}, \\ \dot{\hat{z}}_2 \in L^2 \operatorname{sign}(y - \hat{z}_1), \\ \hat{w}_0(t) = \fint_{[\min\{t,T_1\}, \min\{T_2,t\}]} \hat{z}_2(s) ds \\ \hat{w}_2(t) = \fint_{[\min\{t,T_3\}, \min\{T_4,t\}]} \hat{z}_2(s) ds - \hat{w}_0(t) \\ \hat{S}(t) = \begin{cases} 0 & |\hat{w}_2| \leq \epsilon, \\ \frac{\hat{z}_2(t) - \hat{w}_0}{\hat{w}_2} \mathbf{1}_{t \geq T_4} & \text{otherwise.} \end{cases} \end{cases} \tag{7.44}$$

Remark 7.10. *Note that we consider that $\hat{S}$ is undefined if w_2 is too close to 0. The choice of $\hat{S} = 0$ in that case can be replaced with any value.*

We have the following theorem.

Theorem 7.11. *Let $T_4 > T_3 > T_2 > T_1 > 0$ and $\epsilon, \rho > 0$. Assume Assumptions 7.6 and 7.3. Then there exist thresholds $\mu_1^*, \mu_2^* > 0$, obtained from Corollary 7.3.1 in the first alternative of Assumption 7.3 and from Corollary 7.3.2 in the second alternative, such that for every $\mu_1 > \mu_1^*$ and every $\mu_2 > \mu_2^*$ there exists $L^* > 0$ such that for every $L > L^*$, the corresponding solutions of the coupled system ($\star$)–(7.44) under the open-loop control*

$$u^*(t) = u^0_{\mu_1} \mathbf{1}_{[0,T_2]} + u^1_{\mu_2} \mathbf{1}_{[T_2,T_4]}, \tag{7.45}$$

where u^h_μ is defined in Lemma 7.8, initialized in $B_2(0,\rho) \times B_2(0,\rho)$ satisfy

$$\begin{cases} \hat{w}_0(t) = w_0 & t > T_1, \\ \hat{w}_2(t) = w_2 & t > T_3, \\ \hat{S}(t) = \begin{cases} 0 & |\hat{w}_2| \leq \epsilon, \\ S(x_2(t)) & \textit{otherwise}. \end{cases} & t > T_4. \end{cases} \tag{7.46}$$

Proof. By Corollary 7.3.1 in the first alternative of Assumption 7.3, and by Corollary 7.3.2 in the second alternative, we are assured of the existence of μ_1^* and $\mu_2^* > 0$ such that the control u^* steers $S(x_2)$ into

$$S(x_2(t)) = \begin{cases} 0 & t \in [T_1, T_2], \\ 1 & t \in [T_3, T_4]. \end{cases} \tag{7.47}$$

Since u^* is uniformly bounded and Assumption 7.3 provides a known bound on the induced dynamics of x_2 along the corresponding trajectories, and given the initialization of the observer in the known compact set $B(0, \rho)$, the finite-time convergence of the sliding mode in the prescribed time T_1 ensures the existence of $L^* > 0$ such that for every $L > L^*$,

$$\hat{z}_2 = w_2 S(x_2) + w_0 \quad \text{a.e.} \quad \forall t \geq T_1. \tag{7.48}$$

Combining (7.47) and (7.48), the first averaging window yields $\hat{w}_0(t) = w_0$ for $t > T_1$, while the second yields $⨍_{[T_3,T_4]} \hat{z}_2(s)\, ds = w_0 + w_2$, hence $\hat{w}_2(t) = w_2$ for $t > T_3$. Finally, if $|\hat{w}_2| > \epsilon$, the signal $\hat{S}(t)$ can be reconstructed as $S(x_2(t))$; otherwise, the estimate of $S(x_2(t))$ remains undefined (which is consistent with the inherent ambiguity when the parameter magnitude is too small to distinguish from zero). This completes the proof. □

Remark 7.12. *An alternative adaptive observer for (7.44), given in a more dynamical representation, is*

$$\begin{cases} \dot{\hat{z}}_1 = \hat{z}_2 + L\lceil y - \hat{z}_1 \rfloor^{\frac{1}{2}}, \\ \dot{\hat{z}}_2 \in L^2 \operatorname{sign}(y - \hat{z}_1), \\ \dot{\hat{w}}_0 = \frac{1}{T_2 - T_1} \hat{z}_2(t) 1_{[T_1,T_2]}(t) \\ \dot{\hat{w}}_2 = \frac{1}{T_4 - T_3} (\hat{z}_2(t) - \hat{w}_0(t)) 1_{[T_3,T_4]}(t) \\ \hat{S}(t) = \begin{cases} 0 & |\hat{w}_2| \leq \epsilon, \\ \frac{\hat{z}_2(t) - \hat{w}_0}{\hat{w}_2} \mathbf{1}_{t \geq T_4} & \textit{otherwise}. \end{cases} \end{cases} \tag{7.49}$$

with initialization $\hat{w}_0(0) = \hat{w}_2(0) = 0$. This formulation is motivated by the fact that the average terms in (7.44) (evaluated at the final times T_2 and T_4) are equivalent to the integrals of the above dynamics. Note, however, that the normalization factors in the differential equations are chosen to match the interval lengths of the corresponding averaging windows. The original formulation (7.44) is more general because it computes averages over the intervals $[T_1, t]$ for $t \geq T_1$ and $[T_3, t]$ for $t \geq T_3$, thus providing a valid estimate before the final times T_2 and T_4 respectively.

The next step in the estimation procedure for system ($\star$) is to obtain a valid estimate of the state x_2 after having estimated the parameters w_0, w_2 and the signal $S(x_2)$. To this end, we show in the next section that for systems of the type (7.21), one can design an output-feedback control that takes $S(x_2)$ as output and will steer the state into the invertible region of S. Estimating x_2 is necessary when the unknown function ϕ_2 has additional structure that can be identified and/or to perform an estimation process of successive layers, a topic that will be explored in a subsequent section.

7.3.4 *Output-feedback control*

Just as in Section 7.3.2, consider dynamical systems of the type

$$\begin{cases} \dot{x} = f_0(x,t) + u, \\ y = S(x), \end{cases} \tag{7.50}$$

Assumption 7.13 (Admissible drift for output-feedback)**.** *The function f_0 is continuous and there exists a known constant $m > 0$ such that along the considered trajectories*

$$|f_0(x,t)| \leq m. \tag{7.51}$$

This holds in particular if either $\|f_0\|_\infty \leq m$, or $f_0(x,t) = -\lambda x + \psi(x,t)$ with $\lambda > 0$ and bounded ψ, since in that case (for known bounds on control and initial condition) trajectories remain in a known invariant attracting compact set and $|f_0|$ is bounded there by a known constant m.

We want to design an output-feedback control $u_2^*(t) = \eta(y(t); \mu, \gamma(t), \delta)$ such that in any prescribed time $T > 0$ we estimate x_2 from the knowledge of w_2, w_0 and $S(x_2)$. Define the following sets: $\Omega_S := [\underline{R}, \bar{R}]$ and $\Omega_{S,\epsilon} := [\underline{R} + \epsilon, \bar{R} - \epsilon]$. Let $\gamma \in C^{1,L}(\mathbb{R}, \Omega_{S,\epsilon}), \mu > 0$.

$$\eta(\sigma; \mu, \gamma(t), \delta) = \begin{cases} \mu & \sigma \leq 0, \\ \alpha_1(\sigma; \mu, \gamma) & 0 < \sigma \leq \delta, \\ -\mu \left(S^{-1}(\sigma) - \gamma\right) & \text{if } \delta \leq \sigma \leq 1 - \delta, \\ \alpha_2(\sigma; \mu, \gamma) & 1 - \delta \leq \sigma \leq 1 \\ -\mu & \text{if } \sigma \geq 1. \end{cases} \tag{7.52}$$

Here, α_1, α_2 are affine functions ensuring that η is continuous. Note that the continuity and the introduction of the safety zone indexed by $\delta > 0$ are not strictly necessary, but they simplify the proof, as we will show in the next lemma. Indeed, with η defined as above, for any piecewise continuous function $\tilde{x}$ with a finite number of discontinuities, solutions of system (7.50) exist for all times. This will ease the proof for the coupled plant-observer system.

We have the following lemma.

Lemma 7.14 (Output Feedback Controllability on the set $(\underline{R}, \bar{R})$)**.** *Let* $0 < \delta < 1/2$, $\epsilon > 0$, $T_2 > T_1 > 0$, $\rho > 0$, $m > 0$ *and* $\gamma \in C^{1,L}(\mathbb{R}, \Omega_{S,\delta})$. *Assume Assumption 7.13 holds with this constant m. There exists* $\mu^* > 0$ *such that for every* $\mu > \mu^*$, *solutions of the control system (7.50) under output-feedback control* $u(t) = \eta(y(t); \mu, \gamma(t), \delta)$,

$$\begin{cases} \dot{x} = f_0(x,t) + \eta(y(t); \mu, \gamma(t), \delta), \\ y = S(x), \end{cases} \tag{7.53}$$

with initial condition in $B(0, \rho)$ *exist, are global in time, and verify*

$$x(t) \in [\underline{R} + \delta, \bar{R} - \delta], \quad t \geq T_1. \tag{7.54}$$

Moreover, we have

$$\|x - \gamma\| \to 0. \tag{7.55}$$

and more specifically,

$$\|x - \gamma\| \leq \epsilon, \quad t > T_2. \tag{7.56}$$

Proof. From Assumption 7.2, the function S is continuous. Combined with the continuity of f_0 and η in both x and t, the existence and absolute continuity of solutions follow from standard existence theorems (Peano/Carathéodory). Moreover, the uniform bound on η,

$$\|\eta\| \leq \max\{\mu, 2\mu \max\{\underline{R}, \bar{R}\}\}, \tag{7.57}$$

together with Assumption 7.13, ensures that solutions are globally defined in time (no finite-time escape).

We now prove the first claim: that $x(t)$ enters and remains in the invertible region of S after time T_1. Without loss of generality, we consider the case $x(0) > \bar{R}$ (the case $x(0) < \underline{R}$ is similar). Since S satisfies (7.8), we have $S(x(0)) = 1$.

Let $T_0 > 0$ be arbitrary. By continuity of the solution and using arguments analogous to Lemma 7.8, there exists $\tau > 0$ and an interval $[0, \tau] \subset [0, T_0]$ such that

$$\dot{x} \leq m - \mu. \tag{7.58}$$

Consequently, there exists $\mu_1^* > 0$, depending only on T_0, m, and ρ, such that for any $\mu > \mu_1^*$, we have $x(T_0) \leq \bar{R}$. Next, since α_2 is an affine function, we can choose $\mu_2^* > 0$ such that for all $\mu > \mu_2^*$ and for $x(t) \in [S^{-1}(1-\delta), \bar{R}]$,

$$\dot{x} \leq m + \eta(S(x(t)); \mu, \gamma(t), \delta) \leq m + \min\{-\mu, -\mu(S^{-1}(1-\delta) - (\bar{R} - \delta))\} \leq 0. \tag{7.59}$$

Thus, selecting $\mu > \max\{\mu_1^*, \mu_2^*\}$ ensures the existence of $T_1 > 0$ such that

$$x(t) \in [\underline{R} + \delta, \bar{R} - \delta] \quad \text{for all } t \geq T_1, \tag{7.60}$$

which completes the first part of the proof.

For the second part, we establish exponential convergence toward the reference signal $\gamma(t)$ once $x(t)$ enters the invertible region. Assume that at some $t_0 \geq 0$, we have $S(x(t_0)) \in [\delta, 1-\delta]$. Then, by continuity and the invariance properties established above, $x(t)$ remains in $(\underline{R}, \bar{R})$ for $t \geq t_0$, and almost everywhere we have

$$\dot{x} \leq m - \mu(x - \gamma(t)). \tag{7.61}$$

Since $\gamma \in C^{1,L}$, its derivative satisfies $\|\dot{\gamma}\| \leq L$. Defining $z = x - \gamma$, we obtain

$$2\frac{d}{dt}|z|(t) \leq (m - \dot{\gamma}(t)) - \mu|z|(t). \tag{7.62}$$

Applying Grönwall's inequality, we can choose $\mu_3^* > 0$ such that for any $\mu \geq \mu_3^*$,

$$|z(t)| \to 0 \quad \text{as } t \to \infty, \tag{7.63}$$

and more specifically,

$$|z(t)| \leq \epsilon \quad \text{for all } t \geq t_0 + (T_2 - T_1). \tag{7.64}$$

Taking thus $\mu^* := \max\{\mu_1^*, \mu_2^*, \mu_3^*\}$ completes the proof. □

7.4 ESTIMATOR DESIGN FOR SYSTEM ($\star$)

We have now constructed all the necessary tools to complete the estimation process for system ($\star$), meaning the estimation of w_0, w_2, θ from measurement of x_1. To achieve this, we combine the adaptive observer introduced in (7.44), which achieves finite-time convergence of the parameters and $S(x_2)$, and the output-feedback control (8.35) introduced in the previous section.

We therefore propose the following adaptive observer. Assume $|w_2| > \epsilon$. Let $T_5 > T_4 > T_3 > T_2 > T_1 > 0, L, \epsilon > 0$.

$$\begin{cases} \dot{\hat{z}}_1 = \hat{z}_2 + L\lceil y - \hat{z}_1 \rfloor^{\frac{1}{2}}, \\ \dot{\hat{z}}_2 \in L^2 \operatorname{sign}(y - \hat{z}_1), \\ \hat{w}_0(t) = ⨍_{[\min\{t,T_1\},\min\{T_2,t\}]} \hat{z}_2(s)\, ds \\ \hat{w}_2(t) = ⨍_{[\min\{t,T_3\},\min\{T_4,t\}]} \hat{z}_2(s)\, ds - \hat{w}_0(t) \\ \hat{S}(t) = \begin{cases} 0 & |\hat{w}_2(t)| \leq \epsilon, \\ \frac{\hat{z}_2(t) - \hat{w}_0(t)}{\hat{w}_2(t)} \mathbf{1}_{t \geq T_4} & \text{otherwise}, \end{cases} \\ \hat{x}_2 = S^{-1}(\hat{S}) \mathbf{1}_{t \geq T_5}. \end{cases} \tag{7.65}$$

Let also $(\mu_k)_{0 \leq k \leq 3}$ be a sequence of positive real numbers and define the following output-feedback control:

$$U_2^*(\hat{S}, t; (\mu_k), \gamma) = u_{\mu_1}^0 \mathbf{1}_{[0,T_2]} + u_{\mu_2}^1 \mathbf{1}_{[T_2,T_4]} + \eta(\hat{S}(t); \mu_3, \gamma(t), \delta) \mathbf{1}_{[t \geq T_4]}. \tag{7.66}$$

We have the following theorem.

Theorem 7.15. *Let $T_5 > T_4 > T_3 > T_2 > T_1 > 0$ and $\epsilon, \rho > 0$, $0 < \delta < 1/2$. Assume Assumptions 7.6 and 7.3, and $|w_2| > \epsilon$. Then there exist thresholds $\mu_1^*, \mu_2^* > 0$ such that for every $\mu_1 > \mu_1^*$ and every $\mu_2 > \mu_2^*$ there exists $\mu_3^* > 0$ such that for every $\mu_3 \geq \mu_3^*$ there exists $L^* > 0$ such that for every $L > L^*$, the corresponding solutions of the coupled system ($\star$)–(7.65) under the output-feedback control*

$$u^*(t) = U_2^*(\hat{S}, t; (\mu_k)_k, \gamma) \tag{7.67}$$

initialized in $B_2(0,\rho) \times B_2(0,\rho)$ satisfy

$$\begin{cases} \hat{w}_0(t) = w_0 & t > T_1, \\ \hat{w}_2(t) = w_2 & t > T_3, \\ \hat{S}(t) = S(x_2(t)) & t > T_4, \\ \hat{x}_2(t) = x_2(t) & t > T_5, \\ \|\hat{x}_2 - \gamma\| \to 0. \end{cases} \tag{7.68}$$

Proof. The argument combines Theorem 7.11 and Lemma 7.14. First, Theorem 7.11 guarantees the existence of $\mu_1^*, \mu_2^* > 0$ such that for every $\mu_1 > \mu_1^*$ and $\mu_2 > \mu_2^*$, there exists $L_1^* > 0$ such that for any $L > L_1^*$, solutions exist and we have exact estimates of w_0, w_2, and $S(x_2(t))$ for all $t \geq T_4$.

For any $\mu_3 > \mu_2$, the closed-loop system contains a single time discontinuity (at $t = T_4$) and a spatial discontinuity only through the sign function, which is upper semi-continuous with compact convex values. Furthermore, by Assumption 7.3, ϕ_2 is continuous and bounded on the relevant trajectories, and U_2^* and η are continuous and uniformly bounded; Filippov's existence theorem therefore guarantees global existence of solutions for any $L > 0$. We can then select $L_2^*(\mu_3) > 0$ such that for every $L > L_2^*$ the sliding-mode observer maintains finite-time convergence even under the influence of η, as it is uniformly bounded with a known bound depending on $\mu_3^* > 0$. Setting $L^* := \max\{L_1^*, L_2^*\}$ and taking $L > L^*$ ensures the observer gains are sufficient for both the initial estimation phase and the subsequent output-feedback phase. This ensures that $\hat{S}(t) = S(x_2(t))$ for all times $t \geq T_4$.

Lemma 7.14 then ensures that there exists $\mu_3^* > 0$ such that if μ_3 is chosen above μ_3^*, we have $x_2(t) \in [\underline{R} + \delta, \bar{R} - \delta]$ for all $t \geq T_5$ and $\|x_2 - \gamma\| \to 0$ as $t \to \infty$. Since $\hat{S}(t) = S(x_2(t))$ for all $t \geq T_4$ and S is invertible on $[\underline{R} + \delta, \bar{R} - \delta]$, it follows that $\hat{x}_2(t) = S^{-1}(\hat{S}(t)) = x_2(t)$ for all $t \geq T_5$. Consequently, $\|\hat{x}_2 - \gamma\| \to 0$ as $t \to \infty$. □

8

FINITE DIMENSIONAL SYSTEMS

In this work, we propose a novel method for parameter estimation in highly uncertain nonlinear systems in which the measured output is given by the product of an unknown matrix with a known saturation function evaluated at the (unknown) state. Classical estimation techniques often face significant difficulties in such settings, particularly when the system dynamics are only partially known. Our approach exploits the structural properties of the saturation function itself. Specifically, we construct explicit piecewise-constant control inputs that drive the system state into regions where the saturation output takes predetermined values, without requiring knowledge of the exact state or of the underlying dynamics. This procedure generates a signal originating from an unknown function of the state whose values are known at specified times and whose variation over the considered time interval is sufficiently informative to enable the identification of the unknown parameters. We present a canonical form of dynamical systems for which this procedure can be applied. We further extend these results to show output feedback controllability under slightly stronger condition. The main result is stated in Theorem 8.7. Numerical simulations are given to showcase the method.

8.1 INTRODUCTION

The application of control and estimation methods to nonlinear systems is a fundamental challenge in many areas of science and engineering. Nonlinear phenomena such as saturations or more general activation functions are ubiquitous in real-world systems, and are particularly prevalent in neuroscience. Many widely used models of brain activity (e.g., the Wilson–Cowan model [167, 168] and Amari neural fields [3]), explicitly incorporate such nonlinearities.

All dynamical models depend on a certain number of parameters whose values influence the behavior and the response to controls. It is therefore critical to estimate, exactly or approximately, the value of these parameters. Some of them may be known from physical considerations, but most of the parameters have to be estimated from indirect measurement via *parameter identification* [121]. By identification, we mean any method or process that uses measurements of various outputs of the system corresponding to some inputs (or excitation) to deduce some estimate of the parameters. The choice of the excitation, or experiment design, may be an important part of the process. Moreover, the state of the system may not be fully observed, and the estimation process may have to rely on partial measurements, in this case, the estimation process is called *state and parameter estimation* or *adaptive observation*.

These problems in dynamical systems have been extensively studied since the foundational work in adaptive control and system identification [121, 29]. Classical approaches rely on systems with few parameters to estimate. The standard framework assumes either contraction properties of the system or passivity conditions that ensure convergence of estimation errors [121, 29]. These assumptions are often made to ensure the recovery of the state independently of the unknown parameters. Moreover, it is often assumed that the trajectories of the system verify a persistent excitation condition, which ensure that the system is sufficiently excited to allow for parameter recovery [29]. In particular, persistent excitation of the state trajectories has become the standard sufficient condition to ensure parameter identifiability.

However, these classical frameworks rely on assumptions that do not hold generally. Many physical systems fail to satisfy global contraction or passivity properties. The persistent excitation requirement on unmeasured states is difficult to verify. Additionally, standard methods often require that unknown parameters multiply either a persistent signal, often called the regressor, or a state-dependent function that is assumed persistently excited through the above-mentioned hypothesis.

More and more models of complex systems, such as those in neuroscience, exhibit high uncertainty, with many unknown terms. Usually, only the structure of the model is known. As such, the classical assumptions for parameter estimation are not satisfied, and the standard methods fail to provide a robust method for estimation.

In this chapter, we focus on a class of highly uncertain non-linear systems where the dynamics are unknown and where the output contains an unknown matrix to retrieve applied to the saturation of the state. It is a challenging problem as the unknown parameters do not multiply a known signal, and the system does not satisfy the classical assumptions for estimation [29]. We formalize this class of systems in a canonical form which represent the minimal structural assumptions under which parameter estimation with these techniques is feasible.

We propose a novel approach where we leverage the structure imposed by saturation nonlinearities to design explicit control inputs that steer the system's state into regions where the saturation function takes known values at predetermined values. The estimation procedure consists of three sequential steps. First, linear output parameters are recovered through regression on measurable data with a designed control input. Second, the gain matrix is identified by controlling the system to activate different saturation regions and exploiting the resulting input-output relationship. Third, given complete parameter knowledge, state recovery is achieved through output feedback control combined with algebraic inversion of the saturation function. Each step is supported by explicit control designs and estimation guarantees.

8.2 MAIN RESULT

NOTATIONS

1. For any two integers $p < q$, we denote by $[\![p, q]\!]$ the set of integers $[\![p, q]\!] = \{i \in \mathbb{Z},\ p \leq i \leq q\}$.
2. The vector $\mathbf{1}_d$ denotes the vector of dimension d with 1 whose entries are all equal to 1.

3. Given a function $g : \mathbb{R} \to \mathbb{R}$ and a vector $x \in \mathbb{R}^n$, we denote by $g(x) \in \mathbb{R}^n$ the vector obtained by applying g element-wise, i.e., $g(x) = (g(x_1), \ldots, g(x_n))$.
4. Denote for $L > 0$ and $\Omega \subset \mathbb{R}^n$ by $C^{1,L}(\mathbb{R}, \Omega)$ the set of trajectories taking values in Ω and whose derivative is uniformly bounded by L.
5. For any set $\Omega \subset \mathbb{R}^n$ whose interior is nonempty, we define by $\mathring{\Omega}$ its interior.
6. An *orthant* of $\mathbb{R}^n$ is a region defined by a fixed sign pattern on each coordinate. Throughout this chapter, a generic orthant is denoted $\mathcal{O}$, and $\mathring{\mathcal{O}}$ stands for its interior.

We recall the following

Definition 8.1. *A trajectory $\gamma \in C(\mathbb{R}, \Omega)$, with Ω open subset of $\mathbb{R}^n$, verifies the Persistent Excitation condition (PE) if there exist $a, b, \tau > 0$ such that*

$$aI_n \preccurlyeq \int_{t-\tau}^{t} \gamma(s)\gamma(s)^\top ds \preccurlyeq bI_n. \tag{8.1}$$

We denote the set of PE trajectories taking values in Ω by $\mathrm{PE}(\Omega)$.

Consider the following system

$$\begin{cases} \dot{x} = -\lambda x + A_0 x + \phi(x, y, t) + Bu \\ y = WS(x) + h_0 + G(y, u, t)^\top \theta \end{cases} \tag{8.2}$$

where $y \in \mathbb{R}^{n_y}$, $x \in \mathbb{R}^{n_x}$, $u \in \mathbb{R}^{n_u}$, $\lambda > 0$, $A_0 \in \mathbb{R}^{n_x \times n_x}$, $W \in \mathbb{R}^{n_y \times n_x}$, $\theta \in \mathbb{R}^{n_\theta}$, $B \in \mathbb{R}^{n_x \times n_u}$, $h_0 \in \mathbb{R}^{n_y}$ and the functions $\phi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R} \to \mathbb{R}^{n_x}$, $G : \mathbb{R}^{n_y} \times \mathbb{R}^{n_m} \times \mathbb{R} \to \mathbb{R}^{n_\theta \times n_y}$. For the meaning of each variable, see Table 8.1

Remark 8.2. *The structure of the output y is in its most general form. The terms $h_0 + G^\top(y, u, t)\theta$ are included for generality; the main result is the estimation of W.*

Concerning the nonlinearity S, we make the following assumption.

Assumption 8.3. *The function $S : \mathbb{R} \to \mathbb{R}$ is a known, globally Lipschitz-continuous, piecewise continuously differentiable function. It is characterized by known saturation levels $\underline{R} < \bar{R}$ such that:*

$$x \leq \underline{R} \Rightarrow S(x) = 0 \quad \textit{and} \quad x \geq \bar{R} \Rightarrow S(x) = 1, \tag{8.3}$$

and is strictly increasing on the interval $[\underline{R}, \bar{R}]$. Consequently, $\|S\|_\infty = 1$. For vector arguments $x \in \mathbb{R}^n$, the function is applied component-wise:

$$S(x) = (S(x_1), \ldots, S(x_n))^\top$$

Denote by $R := \max\{|\underline{R}|, |\bar{R}|\}$ the saturation radius. Denote the unsaturated region $\Omega_S := [\underline{R}, \bar{R}]$ and, for $r > 0$, its strict interior $\Omega_{S,r} := [\underline{R} + r, \bar{R} - r]$.

Assumption 8.4. *The dissipation rate $\lambda > 0$ is known. The matrix $A_0 \in \mathbb{R}^{n_x \times n_x}$ is unknown, with $\|A_0\|_\infty \leq \Lambda_0$ for a known constant $\Lambda_0 < \lambda$. Set*

$$\kappa := \frac{\Lambda_0}{\lambda - \Lambda_0}. \tag{8.4}$$

The perturbation ϕ is continuous and satisfies $\|\phi(x, y, t)\|_\infty \leq b_\phi$ for a known constant $b_\phi > 0$.

Assumption 8.5. *There exists $n+1 \le m \le 2^n$ orthants $\mathcal{O}_i$ such that*

$$\forall i \in [\![0, m]\!], \quad \operatorname{Im}(B) \cap \mathring{\mathcal{O}}_i \neq \varnothing. \tag{8.5}$$

Furthermore, one of the orthants is the negative orthant $\mathcal{O}_0 = \{x \in \mathbb{R}^n : x_i \le 0, \forall i\}$. Denote by $v_i = S(z_i)$ with $z_i \in \mathcal{O}_i \cap B(0, R)$ for $i = 1, \ldots, m$.

For each $i \in [\![0, m-1]\!]$, there exists a unit vector $b^i \in \operatorname{Im}(B) \cap \mathring{\mathcal{O}}_i$ such that

$$\kappa < \min_j |b^i_j|. \tag{8.6}$$

We denote $\underline{b}^i := \min_j |b^i_j|$ for such a choice.

Assumption 8.6. *The function G is of class C^1 and verify the following persistence of excitation condition: there exists $\tau_e > 0$, constants $0 < a < b$ such that for any $u \in L^\infty_{loc}(\mathbb{R}_+, \mathbb{R}^{n_u})$ and $y \in L^\infty_{loc}(\mathbb{R}_+, \mathbb{R}^{n_y})$, and for any $t \ge 0$, it holds*

$$a\,\mathrm{Id}_{q_1} \preceq \int_t^{t+\tau_e} \left(G_{u,y}(s) - \bar{G}_{u,y}(s)\right)\left(G_{u,y}(s) - \bar{G}_{u,y}(s)\right)^\top ds \preceq b\,\mathrm{Id}_{q_1}, \tag{8.7}$$

where we defined $G_{u,y}(s) := G(u(s), y(s), s)$, and $\bar{G}_{u,y}(s) := \frac{1}{\tau_e}\int_t^{t+\tau_e} G_{u,y}(\sigma)\, d\sigma$.

We recall all the assumptions and meanings of equation (8.2) in table 8.1

Should the assumptions of Table 8.1 be verified, we have the following theorem that will be proven in Section 8.5.

Theorem 8.7. *Under the assumptions 8.3, 8.5, 8.6, 8.4, for any $\varrho > 0$ and $T, T_0 > 0$ such that $T \ge T_0 + \tau_e$ with τ_e defined in Assumption 8.6, we have the following:*

1. *Let $E_1 : L^\infty([0,T], \mathbb{R}^{n_u}) \times L^\infty([0,T], \mathbb{R}^{n_y}) \to \mathbb{R}^{n_y + n_\theta}$ be defined as*

$$E_1(u, y) = \left(\int_{T_0}^T \Phi(s)\Phi(s)^\top ds\right)^{-1} \int_{T_0}^T \Phi(s) y(s)\, ds, \quad \Phi(t) = \begin{bmatrix} I_{n_y} \\ G_{u,y}(t) \end{bmatrix} \tag{8.13}$$

   Then, there exists constant control $\bar{u}$ such that for any solution of (8.2) initialized in $B(0, \varrho)$ with output $y \in L^\infty([0,T], \mathbb{R}^{n_y})$, it holds

$$E_1(\bar{u}, y) = \begin{bmatrix} h_0 \\ \theta \end{bmatrix}. \tag{8.14}$$

2. *Assume that (θ, h_0) are known. Define $\tilde{y}(t) := y(t) - h_0 - G(y(t), u(t), t)^\top \theta$. Let $t_1, \ldots, t_{n_x+1} \subset [0, T]$ and $\tau > 0$ be such that $t_{i+1} - t_i > \tau$ for all $i = 1, \ldots, n_x$. Let $E_2 : \mathbb{R}^{n_y + n_\theta} \times L^\infty([0,T], \mathbb{R}^{n_u}) \times L^\infty([0,T], \mathbb{R}^{n_y}) \to \mathbb{R}^{n_y \times n_x}$ be defined as*

$$E_2(u, y) = \left[\sum_{i=1}^{n_x} \int_{t_i}^{t_i+\tau} \tilde{y}(s) v_i^\top\, ds\right] \left[\sum_{i=1}^{n_x} \int_{t_i}^{t_i+\tau} v_i^\top v_i\, ds\right]^{-1}, \tag{8.15}$$

   Then, there exists a piecewise constant control $u^\star \in L^\infty([0,T], \mathbb{R}^{n_u})$ such that for any solution of (8.2) initialized in $B(0, \varrho)$ with output $y \in L^\infty([0,T], \mathbb{R}^{n_y})$, it holds

$$E_2(u^\star, y) = W, \tag{8.16}$$

Term	Unknown	To Retrieve?	Meaning	Assumptions
x	✓	Yes	State vector	Evolves according to the system dynamics.
u			Control input	Unconstrained
y			Measured output	Noise free for now.
λ			Dissipation rate	Known constant $\lambda > 0$.
A_0	✓		Coupling matrix	Unknown matrix $A_0 \in \mathbb{R}^{n_x \times n_x}$, with known bound $\Lambda_0 := \lVert A_0 \rVert_\infty$ ($\lVert A_0 x \rVert_\infty \leq \Lambda_0 \lVert x \rVert_\infty$) satisfying $\Lambda_0 < 2\lambda$. Define $\kappa := \Lambda_0 / (\lambda - \Lambda_0)$.
ϕ	✓		Bounded perturbation	Continuous, with $$\sup_{x,y,t} \lVert \phi(x,y,t) \rVert_\infty \leq b_\phi \quad (8.8)$$ for a known constant $b_\phi > 0$.
W	✓	Yes	Unknown parameter matrix	Constant, bounded with known bound; to be estimated.
h_0	✓	Yes	Bias vector	Constant, bounded with known bound; to be estimated.
θ	✓	Yes	Parameter vector	Constant, bounded with known bound; to be estimated.
S			Saturation function	The function $S : \mathbb{R} \to \mathbb{R}$ is a known, globally Lipschitz-continuous, piecewise continuously differentiable function. It is characterized by known saturation levels $\underline{R} < \bar{R}$ such that: $$x \leq \underline{R} \Rightarrow S(x) = 0 \quad \text{and} \quad x \geq \bar{R} \Rightarrow S(x) = 1, \quad (8.9)$$ and is strictly increasing on the interval $[\underline{R}, \bar{R}]$. Consequently, $\lVert S \rVert_\infty = 1$. For vector arguments $x \in \mathbb{R}^n$, the function is applied component-wise: $$S(x) = (S(x_1), \dots, S(x_n))^\top$$ Denote by $R := \max\{\lvert \underline{R} \rvert, \lvert \bar{R} \rvert\}$ the saturation radius.
B			Constant matrix of appropriate dimensions.	The matrix verify the following controllability condition: There exists $n+1 \leq m \leq 2^n$ orthants $\mathcal{O}_i$ such that $$\forall i \in [\![0, m]\!], \quad \mathrm{Im}(B) \cap \mathring{\mathcal{O}}_i \neq \varnothing. \quad (8.10)$$ Furthermore, one of the orthants is the negative orthant $\mathcal{O}_0 = \{x \in \mathbb{R}^n : x_i \leq 0, \forall i\}$. Denote by $v_i = S(z_i)$ with $z_i \in \mathcal{O}_i \cap B(0,R)$ for $i = 1, \dots, m$. For each $i \in [\![0, m-1]\!]$, there exists a unit vector $b^i \in \mathrm{Im}(B) \cap \mathring{\mathcal{O}}_i$ such that $$\kappa < \min_j \lvert b^i_j \rvert. \quad (8.11)$$ We denote $\underline{b}^i := \min_j \lvert b^i_j \rvert$ for such a choice.
G			Regressor function	The function G is of class C^1 and verify the following persistence of excitation condition: there exists $\tau_e > 0$, constants $0 < a < b$ such that for any $u \in L^\infty_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^{n_u})$ and $y \in L^\infty_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^{n_y})$, and for any $t \geq 0$, it holds $$a \,\mathrm{Id}_{q_1} \preceq \int_t^{t+\tau_e} \left(G_{u,y}(s) - \bar{G}_{u,y}(s)\right) \left(G_{u,y}(s) - \bar{G}_{u,y}(s)\right)^\top ds \preceq b \,\mathrm{Id}_{q_1}, \quad (8.12)$$ where we defined $G_{u,y}(s) := G(u(s), y(s), s)$, and $\bar{G}_{u,y}(s) := \frac{1}{\tau_e} \int_t^{t+\tau_e} G_{u,y}(\sigma)\, d\sigma$.

Table 8.1: Summary of variables and assumptions

3. *Assume that $n_u = n$ and that B is invertible. Assume that (W, h_0, θ) are known. Then, for any $\tau > 0$, $L > 0$ and $\varepsilon > 0$, there exists an explicit output feedback law $\eta : \mathbb{R}^{n_y} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_u}$ such that for any target curve $\gamma \in C^{1,L}(\mathbb{R}, \Omega_S)$ with $\|\dot{\gamma}\|_\infty \leq L$, the solution of the system (8.2) under the control $u(t) = \eta(y(t), \gamma(t))$ and initial condition $x(0) \in B(0, \varrho)$ satisfies*

$$\|x(t) - \gamma(t)\| \leq \varepsilon, \quad \forall t \geq \tau \quad \text{and} \quad \lim_{t \to \infty} \|x(t) - \gamma(t)\| = 0. \tag{8.17}$$

In particular, if $\operatorname{dist}(\gamma(t), \Omega_S^c) > \varepsilon$ [1], the state $x(t)$ remains in the unsaturated region for all $t \geq \tau$, and thus it can be recovered by

$$x_i(t) = S^{-1}\left(\left(W^\dagger \tilde{y}(t)\right)_i\right). \tag{8.18}$$

The theorem states that under the assumptions of Table 8.1, we can recover the unknown parameters h_0, θ and W and the state x in a sequential manner. First, should everything be unknown, we can recover h_0 and θ through a simple regression on the output y when the system (8.2) is excited with a specific constant control.

The second point of the theorem states that once h_0 and θ have been recovered or if there were known, we can recover W. One may see the second point of the theorem as the core of the method, as it relies on the design of a specific control input that drives the system into different saturation regions, which allows to generate a signal whose variation is sufficiently rich to recover W.

The last point of the theorem states that once the parameters are recovered, we can essentially force the state, through an output-feedback control, to remain in the unsaturated region, hence allowing to recover the state. Furthermore, we can also steer the state to any desired trajectory in the unsaturated region.

Remark 8.8. *If $h_0 = 0$ then the assumption on B can be relaxed to the existence of n orthants $\mathcal{O}_i$ such that $\operatorname{Im}(B) \cap \mathcal{O}_i \neq 0$ without crossing the negative orthant. We chose to present the theorem with a drift h_0 as it is more relevant to the class of systems that were inspired by [167]. Note however that the presence of h_0 imposes a higher controllability condition on B as it requires the existence of a control direction in the negative orthant.*

Remark 8.9. *The two estimators E_1 and E_2 are explicit and non-asymptotic. The first estimator E_1 is a standard batch least-squares estimator for linear regression, while the second estimator E_2 is a least-squares estimator for the estimation of W through the use of the vectors v_i which serve as input over predetermined time intervals.*

A dynamical reformulation of these estimators can be obtained, which can be of benefit when considering the stability of these estimator. For instance, the first estimator E_1 can be implemented through a Recursive Least-Squares (RLS) algorithm [121], which is a standard adaptive estimation technique in control theory.

Remark 8.10. *In the case where h_0, θ are known or equal to 0, the control of point 2 of Theorem 8.7 can be designed as an output feedback controls. Indeed, the structural assumption on the saturation*

1 Ω_S^c denotes $\mathbb{R}^{n_x} \setminus \Omega_S$

S (Assumption 8.3) ensures that the signal y will be constant when S is saturated and therefore will contain this information. We can thus construct a trigger mechanism to switch the control to a lower value when the system enters a saturation region. This would allow minimize the energy[2] needed to steer the system into the saturation region, and thus to generate the signal required for estimation.

Although the idea is straightforward, the construction of such a control is not trivial. We focus here on presenting the novel ideas and approach rather than the technicalities of the control design, but the construction of such output-feedback control is possible and can be done by using a switching mechanism. This will be the subject of future work.

Remark 8.11. *The invertibility assumption on S in the interval $[\underline{R}, \bar{R}]$ is not required for the point 1 and 2 of Theorem 8.7 as it will become clear in Lemma 8.14 and Lemma 8.17*

Remark 8.12. *The theorem is stated with $S : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ applied component-wise to the full state. If the saturation only affects a proper subset of coordinates, or if $W \in \mathbb{R}^{n_y \times p}$ with $p < n_x$, the three-step procedure carries over verbatim: points 1 and 2 depend only on the image of S and the rank of the resulting v_i-family, while point 3 inverts S on the saturated coordinates only.*

8.3 ON THE SUFFICIENT CONDITION OF THE MATRIX B

Consider now the following set

$$\mathcal{A} := \{S(z) : z \in \bigcup_{0 \le i \le m-1} \left(\mathcal{O}_i \cap \{x \in \mathbb{R}^n : \min_{1 \le j \le n} |x_j| \ge \max\{|\underline{R}|, |\bar{R}|\}\} \right) \}., \tag{8.19}$$

From Assumption 8.3 on the saturation function S, we have that $\mathcal{A} \subset \{0,1\}^n$. The orthants $\mathcal{O}_i$ are defined in Assumption 8.3. The sufficient condition for points 1 and 2 of the Theorem 8.7 to hold is that

$$0 \in \mathcal{A}, \quad \operatorname{rank}(\operatorname{span}(\mathcal{A})) = n \tag{8.20}$$

This will become clear with Lemma 8.14 and Lemma 8.17

The condition on B that its image crosses at least $m \ge n+1$ orthants with the negative one given in Assumption 8.3 ensures then that the above condition is met, as the negative orthant is mapped to zero and each other orthant is mapped to a unique element of $\{0,1\}^n$.

The condition on B is non-trivial, but it is not very restrictive. In particular, it does not require B to be of full rank. For instance, in the case $n = 3$, the matrix B can be of rank 2 and satisfy the assumption by taking two linearly independent columns that belong to different orthants where at least one is in the negative orthant.

In fact, if B is of rank at least $(n+2)/2$, the assumption is verified as long as it crosses the negative orthant. Indeed, any linear subspace of $\mathbb{R}^n$ of rank k intersects at least $2k$ orthants. This gives the minimal number of orthants that a k-dimensional subspace can intersect. In particular, we have the following

2 In the sense of $\|u\|$

Proposition 8.13 (Lower Bound on Orthants Intersected). *If the matrix B is of rank*

$$k = \begin{cases} \frac{n+1}{2}, & \text{if } n \text{ is odd,} \\ \frac{n+2}{2}, & \text{if } n \text{ is even,} \end{cases} \tag{8.21}$$

and contains a column whose entries are either all strictly positive or all strictly negative, then Assumption 8.5 is satisfied.

The maximum number of orthants that a k-dimensional subspace $\mathrm{Im}(B) \subset \mathbb{R}^n$ can intersect is given by the following proposition obtained from hyperplane arrangement theory [169].

Proposition 8.3.1 (Upper Bound on Orthants Intersected). *Let $B \in \mathbb{R}^{n\times k}$ with $\mathrm{rank}(B) = k \leq n$. Then $\mathrm{Im}(B)$ intersects at most $2\sum_{i=0}^{k-1}\binom{n-1}{i}$ orthants of $\mathbb{R}^n$.*

The same result can be obtained by oriented matroid theory, which is related to the number of sign patterns of a linear subspace. See [32, Chapter 4]. The exact condition when the inequality of the above proposition becomes tight are described in [169].

8.4 TECHNICAL LEMMAS FOR THE PROOF OF THE MAIN THEOREM

In this section, we present the technical lemmas that are required to prove the main theorem. These lemmas establish key properties of the system dynamics, the behavior of the saturation function, and the design of control inputs that allows the parameter estimation and state recovery.

Concerning system (8.2), the vector field induced by $x \to -\lambda x + A_0 x + \phi(x, y, t)$ is continuous and linearly bounded. Together with the piecewise continuous control Bu, local existence and absolute continuity of solutions follow from standard theorems (Peano/Carathéodory). Since $\Lambda_0 < \lambda$, the matrix $A := -\lambda I + A_0$ is strictly diagonally dominant with negative diagonal, hence Hurwitz. Solutions are therefore globally defined and satisfy

$$\|x(t)\| \leq R_\mu := \rho + \frac{b_\phi + \mu}{\lambda - \Lambda_0}, \quad \forall t \geq 0, \tag{8.22}$$

for any solution initialized in $B(0, \rho)$ with control satisfying $\|Bu\|_\infty \leq \mu$. This bound is *uniform in time* owing to the dissipative structure of $-\lambda I + A_0$.

The following lemma shows that, for μ large enough, a constant control drives the state permanently into any target saturation region reachable by B, with an explicit formula for the required amplitude.

Lemma 8.14. *Let $\rho > 0$, $T_0 > 0$, and $\mathcal{E} \in \mathcal{A}$. Let $b \in \mathrm{Im}(B)$ be a unit vector in the interior of the orthant associated with $\mathcal{E}$, as provided by Assumption 8.5 and (8.19), and set $\underline{b} := \min_j |b_j|$. Then there exists $\mu^\star > 0$ such that for any $\mu \geq \mu^\star$, the constant control u defined by $Bu = \mu b$ yields, for any solution of (8.2) initialized in $B(0, \rho)$:*

$$S(x(t)) = \mathcal{E} \quad \forall t \geq T_0. \tag{8.23}$$

The sufficient control amplitude is

$$\mu^\star = \frac{1}{\underline{b} - \kappa}\left[\frac{\lambda(R+\rho)}{1 - e^{-\lambda T_0}} + (1+\kappa)\, b_\phi + \Lambda_0 \rho\right]. \tag{8.24}$$

Moreover, the state satisfies the uniform-in-time bound

$$\|x(t)\| \le R_\mu := \rho + \frac{b_\phi + \mu}{\lambda - \Lambda_0}, \quad \forall t \ge 0. \tag{8.25}$$

For $\kappa = 0$ (i.e., $A_0 = 0$), no condition on $\underline{b}$ is needed beyond $\underline{b} > 0$, and the formula reduces to $\mu^\star = (b_\phi + \lambda(R+\rho)/(1-e^{-\lambda T_0}))/\underline{b}$.

Proof. From the assumption on B and the definition of $\mathcal{A}$, we choose the control u such that $Bu = \mu b$ with $b \in \operatorname{Im}(B) \cap \mathcal{O}$, $\|b\| = 1$, and $\mathcal{E} = S(x)$ for some $x \in \mathcal{O}$ with $\|x\| > R$. By the global bound (8.22), $\|x(t)\| \le R_\mu$ for all $t \ge 0$.

Components with $b_i > 0$. The dynamics of the i-th component satisfy

$$\dot{x}_i = -\lambda x_i + (A_0 x)_i + \phi_i(x, y, t) + \mu b_i \ge -\lambda x_i + \mu b_i - \Lambda_0 R_\mu - b_\phi. \tag{8.26}$$

Expanding $\Lambda_0 R_\mu = \Lambda_0 \rho + \frac{\Lambda_0}{\lambda - \Lambda_0}(b_\phi + \mu) = \Lambda_0 \rho + \kappa(b_\phi + \mu)$, the effective forcing is

$$c_\mu^{(i)} := \mu b_i - \Lambda_0 R_\mu - b_\phi = (b_i - \kappa)\mu - (1+\kappa)\, b_\phi - \Lambda_0 \rho. \tag{8.27}$$

Since $\kappa < \underline{b} \le b_i$, we have $c_\mu^{(i)} \to +\infty$ as $\mu \to +\infty$. By the scalar comparison principle applied to the ODE $\dot{w} = -\lambda w + c_\mu^{(i)}$ with $w(0) = -\rho$:

$$x_i(t) \ge -\rho\, e^{-\lambda t} + \frac{c_\mu^{(i)}}{\lambda}(1 - e^{-\lambda t}). \tag{8.28}$$

For $t \ge T_0$, this exceeds $\bar{R}$ provided $c_\mu^{(i)} \ge \lambda(\bar{R} + \rho)/(1 - e^{-\lambda T_0})$. Moreover, since $c_\mu^{(i)}/\lambda > \bar{R}$, the equilibrium of the comparison ODE exceeds $\bar{R}$, so $x_i(t) \ge \bar{R}$ is maintained for all $t \ge T_0$ by the dissipative structure. Hence $S(x_i(t)) = 1$ for all $t \ge T_0$.

Components with $b_i < 0$. The symmetric argument gives

$$\dot{x}_i \le -\lambda x_i + \mu b_i + \Lambda_0 R_\mu + b_\phi = -\lambda x_i - c_\mu^{(i)}, \tag{8.29}$$

where $c_\mu^{(i)} = (|b_i| - \kappa)\mu - (1+\kappa)\, b_\phi - \Lambda_0 \rho > 0$ for μ large enough, and by the comparison principle with $x_i(0) \le \rho$, we obtain $x_i(t) \le \underline{R}$ for all $t \ge T_0$, giving $S(x_i(t)) = 0$.

Taking the worst case over all components (i.e., replacing $|b_i|$ by $\underline{b}$) and solving for μ yields (8.24). The uniform bound (8.25) follows from (8.22). □

Remark 8.15 (On the condition $\kappa < \underline{b}$). *The condition $\kappa < \underline{b}$ ensures that the per-component control amplitude $\mu|b_i|$ exceeds the coupling term $\|A_0\|_\infty R_\mu \sim \kappa\mu$. When $A_0 = 0$ ($\kappa = 0$), this reduces to $\underline{b} > 0$, which holds automatically since b lies in the interior of an orthant. For systems of the form $\dot{x} = -Dx + MS(x) + Bu$ with D a positive diagonal matrix, setting $\lambda = \min_i D_{ii}$, $A_0 = 0$, and $\phi = MS(x)$ with $b_\phi = \|M\|_\infty$ gives $\kappa = 0$, and the lemma applies with no additional condition on κ.*

Remark 8.16 (The general geometric condition on B and extension to general Hurwitz matrices.)**.** *The decomposition $A := -\lambda I + A_0$ with $\Lambda_0 < \lambda$ 8.4 and the condition on B 8.5 are* sufficient *structural assumptions. To understand why, and what is the necessary condition on (A, B) so that theorem 8.7, consider a general Hurwitz matrix A (i.e., all eigenvalues have strictly negative real part). With constant control $Bu = \mu b$, Duhamel's formula gives*

$$x(t) = e^{At}x(0) + \int_0^t e^{A(t-s)}[\phi(s) + \mu b]\, ds = \underbrace{-\mu A^{-1}b}_{=: \mu\tilde{b}} + \underbrace{e^{At}(x(0) + \mu A^{-1}b)}_{r_1(t)\to 0} + \underbrace{\int_0^t e^{A(t-s)}\phi(s)\, ds}_{r_2(t),\ \|r_2\|\le Mb_\phi/\alpha}. \tag{8.30}$$

For large μ, the trajectory converges to the steady-state direction $\mu\tilde{b}$ where $\tilde{b} := -A^{-1}b$, up to the bounded residuals r_1 (transient) and r_2 (perturbation). The saturation $S(z_i(t))$ reaches its extreme value (0 or 1) permanently if and only if each component $\tilde{b}_i$ has the correct sign and is bounded away from zero relative to the residual. The condition on b is therefore:

$$\tilde{b} = -A^{-1}b \in \mathrm{int}(\mathcal{O}), \quad \textit{i.e., } \tilde{b}_i \textit{ has the same sign as } b_i \forall\, i, \tag{8.31}$$

where $\mathcal{O}$ is the target orthant. In other words, the map $b \mapsto -A^{-1}b$ must not "flip" any component's sign, i.e. the image $\mathrm{Im}(-A^{-1}B)$ must cross the same orthants as $\mathrm{Im}(B)$, and the components of $\tilde{b}$ must remain uniformly bounded below.

The decomposition $-\lambda I + A_0$ is a sufficient condition. Indeed, write $A = -\lambda I + A_0$ with $\Lambda_0 := \|A_0\|_\infty < \lambda$. Then

$$-A^{-1} = (\lambda I - A_0)^{-1} = \frac{1}{\lambda}\sum_{k=0}^{\infty}\left(\frac{A_0}{\lambda}\right)^k, \tag{8.32}$$

so $\tilde{b} = -A^{-1}b = \frac{1}{\lambda}b + \frac{1}{\lambda}\sum_{k=1}^{\infty}(A_0/\lambda)^k b$. The leading term b/λ is in the same orthant as b, and the perturbation is bounded by $\|\tilde{b} - b/\lambda\|_\infty \le \kappa/\lambda$ where $\kappa = \Lambda_0/(\lambda - \Lambda_0)$. Consequently, $\tilde{b}_i$ has the same sign as b_i whenever $\kappa < |b_i|/\lambda \cdot \lambda = |b_i|$, i.e., whenever $\kappa < \underline{b}$. This is precisely the condition appearing in Lemma 8.14. Moreover, the decomposition gives the explicit constants $M = 1$ and $\alpha = \lambda - \Lambda_0$ (since $e^{At} = e^{-\lambda t}e^{A_0 t}$ and $\|e^{A_0 t}\| \le e^{\Lambda_0 t}$), which is what enables the constructive $\mu^\star$ formula (8.24) and the component-wise comparison proof.

For a general Hurwitz A, the condition (8.31) may or may not hold depending on the unknown entries of A, and it cannot be verified from the known quantities alone, this is the most important restriction. Even if $\mathrm{Im}(B)$ crosses the right orthants, $\mathrm{Im}(-A^{-1}B)$ may not, because the off-diagonal entries of A^{-1} can flip signs. We discuss the sufficient condition on the controllability and the extension to general Hurwitz matrices when the linear part of the system is known in more details in Section 8.3. In summary, the condition on B stated in Table 8.1 is a checkable *sufficient condition that, together with $\kappa < \underline{b}$, guarantees the true geometric condition (8.31). An extension to general Hurwitz matrices is possible provided one assumes additionally that M, α, and the sign pattern of $-A^{-1}B$ are known, but this requires significantly more information about the system.*

The following lemma extends this to a sequence of successive saturation targets over disjoint time intervals.

Lemma 8.17. *Let $T > 0$, $m \in \mathbb{N}$, and $t_1, \cdots, t_{m+1} \subset [0, T]$ and $\tau > 0$ be such that $t_{i+1} - t_i > \tau$ for all $i = 0, \ldots, m+1$ with $t_0 := 0$ and $t_{m+2} := T$. For any sequence of vectors $(v_i)_{0\le i\le m} \subset \mathcal{A}$, there* *exists a piecewise constant control* $u^\star \in L^\infty([0,T],\mathbb{R}^{n_u})$ *such that for any solution of* (8.2) *initialized in* $B(0,\rho)$*:*

$$S(x(t)) = v_i \quad \forall t \in [t_i, t_i+\tau], \quad 1 \le i \le m. \tag{8.33}$$

Proof. For each j, choose $Bu_j = \mu_j b^j$ with $b^j \in \mathrm{Im}(B) \cap \mathcal{O}_j$, $\|b^j\| = 1$, and $\mu_j > 0$. The proof proceeds by induction.

Initial step ($k=1$): Apply Lemma 8.14 with $T_0 = t_1$ and initial radius $\rho_1 := \rho$. There exists $\mu_1 \ge \mu_1^\star$ such that the constant control $u^\star(t) = u_1$ for $t \in [0, t_1+\tau]$ yields $S(x(t)) = v_1$ for all $t \in [t_1, t_1+\tau]$. By the uniform bound (8.25), the state satisfies $\|x(t)\| \le \rho_2 := \rho + (b_\phi + \mu_1)/(\lambda - \Lambda_0)$ for all $t \in [0, t_1+\tau]$.

Induction step: Suppose that after the $(k-1)$-th step, $\|x(t_{k-1}+\tau)\|_\infty \le \rho_k$. Apply Lemma 8.14 on the interval $[t_{k-1}+\tau,\, t_k+\tau]$ with initial radius ρ_k and initial time $t_{k-1}+\tau$ and $T_0 = t_k - t_{k-1} - \tau > 0$. There exists $\mu_k \ge \mu_k^\star(\rho_k, T_0)$ such that $u^\star(t) = u_k$ for $t \in [t_{k-1}+\tau, t_k+\tau]$ yields $S(x(t)) = v_k$ for all $t \in [t_k, t_k+\tau]$. The updated radius bound is

$$\rho_{k+1} = \rho_k + \frac{b_\phi + \mu_k}{\lambda - \Lambda_0}. \tag{8.34}$$

Iterating up to $k = m$ and assembling the piecewise constant control concludes the proof. □

Let $\gamma \in C^{1,L}(\mathbb{R}, \Omega_{S,\delta})$, $\mu > 0$ with $\Omega_{S,\delta}$ defined in Assumption 8.3. We define the following

$$\eta(\sigma;\mu,\gamma,\delta) = \begin{cases} \mu & \sigma \le \ 0, \\ \alpha_1(\sigma;\mu,\gamma) & 0 < \sigma \le \delta, \\ -\mu\left(S^{-1}(\sigma) - \gamma\right) & \text{if } \delta \le \sigma \le 1-\delta, \\ \alpha_2(\sigma;\mu,\gamma) & 1-\delta \le \sigma \le 1 \\ -\mu & \text{if } \sigma \ge 1, \end{cases} \tag{8.35}$$

where α_1 and α_2 are affine functions ensuring the continuity of η.

We have the following lemma, which states that the system is output feedback locally controllable to the unsaturated region Ω_S in arbitrarily small time. This result is what allows recovering the state z from the output y once the parameters W, h_0, θ are estimated and the state is in the unsaturated region.

Lemma 8.18. *Given* W*. Let* $0 < \delta < 1/2$, $\epsilon > 0$, $T_2 > T_1 > 0$, $\rho > 0$, $\gamma \in C^{1,L}(\mathbb{R}, \Omega_{S,\delta})$, $a, b > 0$*. There exists* $\mu^* > 0$ *such that for every* $\mu > \mu^*$ *solutions of the control system*

$$\begin{cases} \dot{z} = f(z) + Bu, \\ \tilde{y}(t) = WS(z(t)), \end{cases} \tag{8.36}$$

with f *continuous function satisfying* $\|f(z)\| \le a\|z\| + b$ *for all* z*, and under the control*

$$u(t) = B^{-1}\eta(y_s(t);\mu,\gamma(t),\delta), \tag{8.37}$$

with $y_S(t) = W^\dagger \tilde{y}(t) = S(z(t))$ *and initial condition in* $B(0,\rho)$ *exist, are global in time, and verify*

$$z(t) \in [\underline{R}+\delta, \bar{R}-\delta]^{n_z}, \quad t \ge T_1. \tag{8.38}$$

Moreover, we have

$$\|z - \gamma\| \to 0. \tag{8.39}$$

and more specifically,

$$\|z - \gamma\| \leq \epsilon, \quad t > T_2. \tag{8.40}$$

Above, $W^{\dagger}$ represents the left-inverse of W. It is guaranteed through the assumption in Table 8.1

Proof. From the continuity of S, f, and η in z, y, t, the existence and absolute continuity of solutions follow from standard existence theorems (Peano/Carathéodory). Moreover, the uniform bound on η,

$$\|\eta\| \leq \max\{\mu, 2\mu \max\{|\underline{R}|, |\bar{R}|\}\}, \tag{8.41}$$

together with the stated growth bound of f ensures that solutions are globally defined in time (no finite-time escape).

Without loss of generality, we consider the and $z \in \mathbb{R}$. This is due to the fact that the feedback η acts independently on each component of z and S. Similarly, we restrcit ourselves to the case $z(0) > \bar{R}$ since the case $z(0) < \underline{R}$ is symmetric.

We now prove the first claim: that $z(t)$ enters and remains in the invertible region of S after time T_1. Since S satisfies Assumption 8.3, we have $S(z(0)) = 1$. When $S(z) = 1$, the feedback gives $\eta = -\mu$, and the scalar dynamics satisfy

$$\dot{z} \leq a|z| + b - \mu. \tag{8.42}$$

Since $z > \bar{R} > 0$, we have $|z| = z$ and thus $\dot{z} \leq a \max\{\rho, R\} + b - \mu$. Using an ode comparison argument, there exists $\mu_1^* > 0$, depending only on T_1, a, b, and ρ, such that for any $\mu > \mu_1^*$, we have $z(T_1) \leq \bar{R}$. Next, since α_2 is an affine function, we can choose $\mu_2^* > 0$ such that for all $\mu > \mu_2^*$ and for $z(t) \in [S^{-1}(1-\delta), \bar{R}]$,

$$\dot{z} \leq a\bar{R} + b + \eta(S(z), t) \leq 0. \tag{8.43}$$

Thus, selecting $\mu > \max\{\mu_1^*, \mu_2^*\}$ ensures the existence of $T_1 > 0$ such that

$$z(t) \in [\underline{R} + \delta, \bar{R} - \delta] \quad \text{for all } t \geq T_1, \tag{8.44}$$

which completes the first part of the proof.

For the second part, we establish exponential convergence toward the reference signal $\gamma(t)$ once $z(t)$ enters the invertible region. Assume that at some $t_0 \geq 0$, we have $S(z(t_0)) \in [\delta, 1-\delta]$. Then, by continuity and the invariance properties established above, $z(t)$ remains in $(\underline{R}, \bar{R})$ for $t \geq t_0$, and almost everywhere we have

$$\dot{z} \leq a \max\{\underline{R}, \bar{R}\} + b - \mu(z - \gamma(t)). \tag{8.45}$$

Since $\gamma \in C^{1,L}$, its derivative satisfies $\|\dot{\gamma}\| \leq L$. Defining $e = z - \gamma$, we obtain

$$\frac{d}{dt}|e|(t) \leq (a \max\{\underline{R}, \bar{R}\} + b + L) - \mu|e|(t). \tag{8.46}$$

Applying Grönwall's inequality, we can choose $\mu_3^* > 0$ such that for any $\mu \geq \mu_3^*$,

$$|e(t)| \to 0 \quad \text{as } t \to \infty, \tag{8.47}$$

and more specifically,

$$|e(t)| \leq \epsilon \quad \text{for all } t \geq t_0 + (T_2 - T_1). \tag{8.48}$$

Taking thus $\mu^* := \max\{\mu_1^*, \mu_2^*, \mu_3^*\}$ completes the proof. □

8.5 PROOF OF THEOREM 8.7

In this section, we present the proof of Theorem 8.7. The proof is structured in three main steps, corresponding to the three statements of the theorem.

Proof part one of Theorem 8.7. From Lemma 8.14 and the assumptions on S, B (Assumptions 8.3–8.5), the negative orthant intersects the image of B and therefore from the construction of the set $\mathcal{A}$ (8.19), we have that $0 \in \mathcal{A}$. We can then apply Lemma 8.14 with $\mathcal{E} = 0$ to find a constant control $\bar{u}$ such that for any solution of (8.2) initialized in $B(0, \varrho)$ with output $y \in L^\infty([0, T], \mathbb{R}^{n_y})$, it holds that $S(x(t)) = 0$ for all $t \in [T_0, T]$. This implies that for all $t \in [T_0, T]$

$$y(t) = h_0 + G(y(t), u(t), t)^\top \theta. \tag{8.49}$$

The model can be rewritten as a regression that is linear in the unknown parameters. Define the augmented parameter vector

$$\zeta := \begin{bmatrix} h_0 \\ \theta \end{bmatrix} \tag{8.50}$$

and the augmented regressor

$$\Phi(t) := \begin{bmatrix} I & G(y(t), u(t), t)^\top \end{bmatrix}. \tag{8.51}$$

Then the system can be written as

$$y(t) = \Phi(t)\zeta. \tag{8.52}$$

This is a linear regression model in the unknown parameters ζ. If the augmented regressor $\Phi(t)$ is persistently exciting as it is ensured by the condition on G given in Assumption 8.6 [121] then the least-squares estimator E_1 defined in the statement of the theorem indeed verify that $E_1(\bar{u}, y) = \zeta$ as $T - T_0 > \tau_e$. This concludes the proof of the first point of the theorem. □

We focus now on the second point of the theorem which obtained as a corollary of Lemma 8.17.

Proof part two of Theorem 8.7. Recall auxiliary output $\tilde{y}(t) := y(t) - h_0 - G(y(t), u(t), t)^\top \theta$. From Lemma 8.17 and the assumptions on S, B detailed in Assumptions 8.3–8.5, the image of B intersect with at least n orthant which are not the negative one. From the construction of the set $\mathcal{A}$ we deduce that there exists a family of vectors $(v_i)_{1 \leq i \leq n}$ in $\mathcal{A}$ that spans the space $\mathbb{R}^n$ and therefore the matrix $M := \sum_{i=1}^n \int_{t_i}^{t_i+\tau} v_i^\top v_i \, ds$ is invertible. We can then apply Lemma 8.17

to find a sequence of constant controls $(u_i)_{1\leq i\leq n}$ such that for any solution of (8.2) initialized in $B(0,\varrho)$ with output $y \in L^\infty([0,T],\mathbb{R}^{n_y})$, it holds that $S(x(t)) = v_i$ for all $t \in [t_i, t_i+\tau]$ and all $i = 1,\dots,n$. This implies that for all $i = 1,\dots,n$ and all $t \in [t_i, t_i+\tau]$ we have that

$$y(t) = Wv_i + h_0 + G(y(t),u(t),t)^\top\theta, \quad t \in [t_i, t_i+\tau], \tag{8.53}$$

and therefore

$$\tilde{y}(t) = Wv_i, \quad t \in [t_i, t_i+\tau]. \tag{8.54}$$

The estimator E_2 defined in the statement of the theorem then verifies that $E_2(u,y) = W$.

□

Part three of the theorem is obtained as a corollary of Lemma 8.18.

Remark 8.19. *The estimator E_2 can be seen as the discret version of the analog estimator in continuous part which was estimator E_1. The estimator E_1 is a least-squares estimator that relies on the persistent excitation of the regressor $\Phi(t)$, while the estimator E_2 is a simple algebraic estimator that relies on the invertibility of the matrix M constructed from the vectors v_i. The control design in Lemma 8.17 ensures that $S(x(\cdot))$ equals the predetermined vectors v_i over the prescribed time intervals.*

8.6 THE CASE OF A KNOWN LINEAR PART

In the preceding sections, the linear part of the dynamics was decomposed as $A = -\lambda I + A_0$ with λ known and A_0 unknown. Let us now examine the case where A is more general but known. Consider then the system

$$\begin{cases} \dot{x} = Ax + \phi(x,y,t) + Bu, \\ y = WS(x) + h_0 + G(y,u,t)^\top\theta, \end{cases} \tag{8.55}$$

where $A \in \mathbb{R}^{n_x\times n_x}$ is *known* and Hurwitz, ϕ is an unknown bounded perturbation with $\|\phi\|_\infty \leq b_\phi$, and all other terms are as in Table 8.1. Since A is Hurwitz, A^{-1} exists and there exist constants $M \geq 1$, $\alpha > 0$ such that $\|e^{At}\| \leq Me^{-\alpha t}$ for all $t \geq 0$.

Using Duhamel's formula with a constant control $Bu = \mu b$ for some $b \in \mathrm{Im}(B)$ we obtain

$$x(t) = \underbrace{-\mu A^{-1}b}_{=:\,\mu\tilde{b}} + \underbrace{e^{At}(x(0)+\mu A^{-1}b)}_{r_1(t)} + \underbrace{\int_0^t e^{A(t-s)}\phi(s)\,ds}_{r_2(t)}, \tag{8.56}$$

where $\tilde{b} := -A^{-1}b$. Since A is known, the vector $\tilde{b} = -A^{-1}b$ is computable for each steering direction $b \in \mathrm{Im}(B)$. The steering condition is therefore checked directly on $\mathrm{Im}(-A^{-1}B)$: we require that there exits at least $n+1 \leq m \leq 2^n$ vectors $\tilde{b}^j := (-A^{-1}b)^j$ that generate, through the orthants $\mathcal{O}_j$ they belong to, sufficiently many distinct saturation vectors spanning $\mathbb{R}^{n_x}$. This replaces the condition $\kappa < \underline{b}^j$, which was a sufficient condition ensuring that $\tilde{b}^j$ stays in the same orthant as b^j when A was unknown. Formally, the assumption on A is as follows:

Assumption 8.20. *The matrix A is known and Hurwitz, and the set*

$$\mathcal{A}' := \{S(-A^{-1}Bx) : x \in \mathbb{R}^m, |x_i| \geq R = \max\{\bar{R}, \underline{R}\}\} \tag{8.57}$$

contains 0 and at least n vectors that span $\mathbb{R}^n$.

For more details on the meaning of this condition and how to check it, see the discussion in Section 8.3. We consider now two subcases, a general case where only the above assumption hold and one where we have controlability of the pair (A, B)

Remark 8.21. *The condition $\kappa < \underline{b}^j$ of Assumption 8.5 was precisely a sufficient condition ensuring $-A^{-1}b^j \in \operatorname{int}(\mathcal{O}_j)$ when A is unknown, where the $\mathcal{O}_j$ are defined in Assumption 8.5.*

When A is known, one computes $-A^{-1}B$ directly. If any $\tilde{b}^j$ has a zero component, the corresponding orthant is not reachable; this is an intrinsic limitation of the pair (A, B) that cannot be circumvented.

Remark 8.22 (General saturation functions)**.** *The same argument extends to cone-valued saturation functions; see Section 8.7 for the precise statement under Assumption 8.25.*

8.6.1 *Known linear part: General Case*

We have the following result, which is the counterpart of Lemma 8.14 in the case of a known matrix A:

Lemma 8.23. *Under assumption 8.20, let $\rho > 0$ and $\mathcal{E} \in \mathcal{A}'$. Let $b \in \operatorname{Im}(B)$ with $\|b\| = 1$, such that for $\tilde{b} := -A^{-1}b$ with $S(l\tilde{b}) = \mathcal{E}$ for a large enough $l > 0$. Assume that the i-th entry$\tilde{b}_i \neq 0$ of $\tilde{b}$ is nonzero for all i, and let $\underline{\tilde{b}} := \min_i |\tilde{b}_i| > 0$. Define the minimal steering time*

$$T_{\min} := \frac{1}{\alpha} \ln\left(\frac{M\|\tilde{b}\|_\infty}{\underline{\tilde{b}}}\right). \tag{8.58}$$

Then for any $T_0 > T_{\min}$, there exists $\mu^\star > 0$ such that for any $\mu \geq \mu^\star$, the constant control $Bu = \mu b$ yields, for any solution of (8.55) initialized in $B(0, \rho)$:

$$S(x(t)) = \mathcal{E} \quad \forall t \geq T_0. \tag{8.59}$$

Setting $\gamma(T_0) := \underline{\tilde{b}} - Me^{-\alpha T_0}\|\tilde{b}\|_\infty > 0$, the sufficient control amplitude is

$$\mu^\star = \frac{R + M\rho + Mb_\phi/\alpha}{\gamma(T_0)}, \tag{8.60}$$

where R is the saturation level 8.5. Moreover, the state satisfies the uniform bound

$$\|x(t)\| \leq \mu\|\tilde{b}\|(1 + M) + M\rho + \frac{Mb_\phi}{\alpha} =: R_\mu, \quad \forall t \geq 0. \tag{8.61}$$

Proof. Rewriting (8.56) by grouping the μ-dependent terms:

$$x(t) = \mu(I - e^{At})\tilde{b} + e^{At}x(0) + \int_0^t e^{A(t-s)}\phi(s)\, ds. \tag{8.62}$$

Consider the case $\tilde{b}_i > 0$. At component i, for $t \geq T_0$:

$$x_i(t) \geq \mu\big(\tilde{b}_i - Me^{-\alpha T_0}\|\tilde{b}\|_\infty\big) - M\rho - \frac{Mb_\phi}{\alpha} = \mu\,\gamma(T_0) - M\rho - \frac{Mb_\phi}{\alpha}, \tag{8.63}$$

using $|((I - e^{At})\tilde{b})_i| \geq \tilde{b}_i - \|e^{At}\|_\infty\|\tilde{b}\|_\infty \geq \tilde{b}_i - Me^{-\alpha T_0}\|\tilde{b}\|_\infty$ for $t \geq T_0$ and $|(e^{At}x(0))_i| \leq M\rho$. The coefficient $\gamma(T_0) > 0$ precisely when $T_0 > T_{\min}$. Then $x_i(t) \geq R$ for $\mu \geq \mu^\star$. The symmetric argument gives $x_i(t) \leq -R$ when $\tilde{b}_i < 0$. Since $e^{-\alpha t}$ only decreases for $t > T_0$, the saturation is maintained permanently. The state bound (8.61) follows from the triangle inequality. $\square$

From this lemma, following the same sequential steering/controllability structure, one can obtain the same results as Theorem 8.7 with the same estimator design, but with the condition on B replaced by the condition on $\mathrm{Im}(-A^{-1}B)$ in Assumption 8.20.

8.6.2 *Known linear part: Controllable Case*

The minimal time $T_{\min}$ in Lemma 8.23 arises because a constant control produces a transient that grows linearly in μ. This can be eliminated under a controllability assumption. Recall that the pair (A, B) is controllable if $\mathrm{rank}[B, AB, \dots, A^{n-1}B] = n$. We recall also that the controllability Gramian

$$W_B(T) := \int_0^T e^{As}BB^\top e^{A^\top s}\,ds \tag{8.64}$$

is positive definite for every $T > 0$ under such condition. Recalling known results from on controllability of linear dynamical systems (One may see the books [158, 55, 148] and the references therein for a general theory), the linear map $u(\cdot) \mapsto \int_0^T e^{A(T-s)}Bu(s)\,ds$ from $L^2([0,T];\mathbb{R}^m)$ to $\mathbb{R}^n$ is surjective: for any target $v \in \mathbb{R}^n$, the open-loop control $u(t) = B^\top e^{A^\top(T-t)}W_B(T)^{-1}v$ achieves $\int_0^T e^{A(T-s)}Bu(s)\,ds = v$ exactly.

Lemma 8.24. *Under the assumptions of Lemma 8.23, assume additionally that (A, B) is controllable. Then for any $T_0 > 0$, there exists $\mu^\star > 0$ such that for $\mu \geq \mu^\star$ there exists a bounded time-varying control $u_\mu : [0, +\infty) \to \mathbb{R}^m$ such that for any solution of (8.55) initialized in $B(0, \rho)$:*

$$S(x(t)) = \mathcal{E} \quad \forall t \geq T_0. \tag{8.65}$$

Proof. Define the control

$$u(t) = \begin{cases} \mu\,B^\top e^{A^\top(T_0-t)}W_B(T_0)^{-1}\tilde{b}, & t \in [0, T_0], \\ u_{\mathrm{cst}}, & t > T_0, \end{cases} \tag{8.66}$$

where $u_{\mathrm{cst}} \in \mathbb{R}^m$ is chosen so that $Bu_{\mathrm{cst}} = \mu b$. On $[0, T_0]$, the Gramian property gives $\int_0^{T_0} e^{A(T_0-s)}Bu(s)\,ds = \mu\tilde{b}$. By Duhamel's formula:

$$x(T_0) = \mu\tilde{b} + r, \qquad \|r\| \leq M\rho + \frac{Mb_\phi}{\alpha}, \tag{8.67}$$

where $r := e^{AT_0}x(0) + \int_0^{T_0} e^{A(T_0-s)}\phi(s)\,ds$ is independent of μ. For $t \geq T_0$, Duhamel's formula with constant control $Bu_{\text{cst}} = \mu b$ and initial condition $x(T_0)$ gives

$$x(t) = \mu\tilde{b} + e^{A(t-T_0)}r + \int_{T_0}^{t} e^{A(t-s)}\phi(s)\,ds, \tag{8.68}$$

since $x(T_0) + \mu A^{-1}b = r$. Both residual terms are bounded independently of μ, so $x_i(t)$ has the sign of $\tilde{b}_i$ for μ large enough. Hence $S(x(t)) = \mathcal{E}$ for all $t \geq T_0$ with $\mu^\star = (R + M\rho + Mb_\phi/\alpha)/\underline{\tilde{b}}$. □

8.6.3 *Conclusion*

The two lemmas above are the analogues of Lemma 8.14 for a general known Hurwitz matrix A. The condition on steering the saturations functions is now imposed on $\text{Im}(-A^{-1}B)$ instead of $\text{Im}(B)$, and since A is known, this condition is directly checkable. Moreover, one can also verify whether the pair (A, B) is controllable. Under these two conditions, the rest of the estimation strategy, mainly, the sequential saturation protocol as in Lemma 8.17, parameter recovery from the output equations, and the main theorem (Theorem 8.7) applies verbatim to system (8.55).

8.7 EXTENSION TO GENERALIZED SATURATION FUNCTIONS

The orthant geometry of Chapter 8 is specific to component-wise saturations. We show here that the estimation strategy extends to a more general class of saturation functions, characterized as follows.

Assumption 8.25. *The function $f : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ is Lipschitz and piecewise C^1. There exist $m+1 \geq n_x$ closed cones $K_0, K_1, \ldots, K_m \subset \mathbb{R}^{n_x}$ with non-empty interior, vectors $v_1, \ldots, v_m \in \mathbb{R}^{n_x}$ generating $\mathbb{R}^{n_x}$, and $R > 0$ such that*

$$f(x) = 0, \qquad \forall x \in K_0,\ \|x\| \geq R, \tag{8.69}$$
$$f(x) = v_i, \qquad \forall x \in K_i,\ \|x\| \geq R, \quad i = 1, \ldots, m. \tag{8.70}$$

The input matrix B satisfies $\text{Im}(B) \cap \mathring{K}_i \neq \emptyset$ for all $i \in [\![0, m]\!]$.

A simple example is $f(x) := S(Hx)$ for some matrix $H \in \mathbb{R}^{n_x \times n_x}$, where S is the component-wise saturation of the preceding sections; taking $H = I$ recovers the orthant case. In general, the cones K_i need not be preimages of orthants and f need not be of this form.

The following lemma is the counterpart of Lemma 8.14 in this cone setting.

Lemma 8.26. *Let Assumption 8.25 hold and let system (8.2) be given. Let $\rho > 0$, $T_0 > 0$, and fix $i \in [\![1, m]\!]$. Let $b^i \in \text{Im}(B) \cap \mathring{K}_i$ with $\|b^i\| = 1$, and let $\delta_i > 0$ denote the angular margin of b^i inside K_i, i.e., the angle from b^i to the nearest face of K_i. Assume*

$$\kappa < \sin(\delta_i). \tag{8.71}$$

Then there exists $\mu_i^* > 0$ such that for any $\mu \geq \mu_i^*$, the constant control $Bu = \mu b^i$ yields, for any solution of (8.2) *initialized in $B(0, \rho)$:*

$$f(x(t)) = v_i, \quad \forall t \geq T_0. \tag{8.72}$$

Moreover, the state satisfies the uniform-in-time bound (8.22).

Remark 8.27. *When K_i is an orthant, the angular margin satisfies $\sin(\delta_i) = \underline{b}^i := \min_j |b_j^i|$, so condition* (8.71) *reduces to $\kappa < \underline{b}^i$, recovering Lemma 8.14 exactly. The obstruction is the same as in Section 8.3: should $\kappa \geq \sin(\delta_i)$, the coupling A_0 can rotate $-A^{-1}b^i$ out of K_i and the cone cannot be reached by a constant control.*

The sequential steering procedure of Lemma 8.17 then extends verbatim using Lemma 8.26: one steers the system successively into $K_{i_1}, K_{i_2}, \ldots$, and the estimator

$$E(u^\star, y) = \left[\sum_{i=1}^{m} \int_{t_i}^{t_i+\tau} \tilde{y}(s)\, v_i^\top \, ds\right] \left[\tau \sum_{i=1}^{m} v_i v_i^\top\right]^{-1} \tag{8.73}$$

recovers W exactly, since $v_1, \ldots, v_m$ are linearly independent by Assumption 8.25. The uniform-in-time state bound (8.22) carries over unchanged, keeping the sequential radius bounds linear. Theorem 8.7 therefore holds with orthants replaced by cones and $\mathcal{A}$ replaced by $\{v_1, \ldots, v_m\}$. The output-feedback point (3) extends under the additional assumption that f is diffeomorphic to a component-wise saturation via a known map h, so that the scalar law η (8.35) applies component-wise.

8.8 NUMERICAL SIMULATION

In this section we present a numerical simulation to illustrate the results of Theorem 8.7. We focus on point 2 of the theorem, namely the estimation of the matrix W, as this aspect is particularly transparent in dimension 2. We assume here $h_0 = 0$ and $\theta = 0$. We set oursleves in dimension $n_x = 2$ and, and we consider the system

$$\begin{cases} \dot{x} = -\lambda x + MS(x) + w_k(t) + Bu(t) \\ y = WS(x) + \nu_k(t) \end{cases} \tag{8.74}$$

where $B = \begin{bmatrix}1 & -1\end{bmatrix}^\top$, M and W are unknown matrices that have been drawn uniformly, $\lambda = 1.2$, and w_k, ν_k are independent zero-mean Gaussian noise processes with standard deviations $\sigma_w = 0.5$ and $\sigma_\nu = 0.3$, respectively. The saturation function S is piecewise-linear with levels $\underline{R} = 0$ and $\bar{R} = 1.5$:

$$S(x) = \mathrm{clamp}\Big(\frac{x}{1.5}, 0, 1\Big) = \begin{cases} 0 & x \leq 0, \\ \dfrac{x}{1.5} & 0 < x < 1.5, \\ 1 & x \geq 1.5. \end{cases} \tag{8.75}$$

Here, we have set $A_0 = 0$.

According to Theorem 8.7, we chose $\tau = 10$, $t_1 = \tau$, $t_2 = t_1 + 2\tau$, $t_3 = t_2 + 2\tau$. The control was chosen according to Lemma 8.17, with $u_1(t) = \mu_1 \mathbf{1}_{[0,t_1+\tau]}(t)B$, and $u_2(t) = -\mu_2 \mathbf{1}_{[t_1+\tau,t_2+\tau]}(t)B$ with $\mu_1 = 10$ and $\mu_2 = 12$.

The controllability condition is verified by Remark 8.8: the image of B intersects two non-negative orthants, which is sufficient. In this case we have $\mathcal{A} = \{(1,0)^\top, (0,1)^\top\}$, which spans $\mathbb{R}^2$, so the invertibility condition on M in Theorem 8.7 is satisfied.

The matrix W is then estimated via the algebraic estimator E_2 defined in Theorem 8.7. The estimation error $\|E_2(u,y) - W\|$, averaged over 100 simulations, was 9×10^{-3}. This small error is a consequence of two complementary noise-rejection mechanisms: first, the saturation function S renders the estimator insensitive to state noise w_k during the saturated phases, since $S(x(t)) = v_i$ independently of small perturbations in x; second, since the output noise ν_k is zero-mean, the time-averaging in E_2 acts as a low-pass filter, reducing the noise contribution by a factor of order $\tau^{-1/2}$.

To visualize the result, Figure 8.1 shows the trajectory x in the phase plane, with the segments where $S(x(t)) = v_i$ for $i = 1, 2$ highlighted. These two saturated segments are the ones exploited by E_2 to recover W.

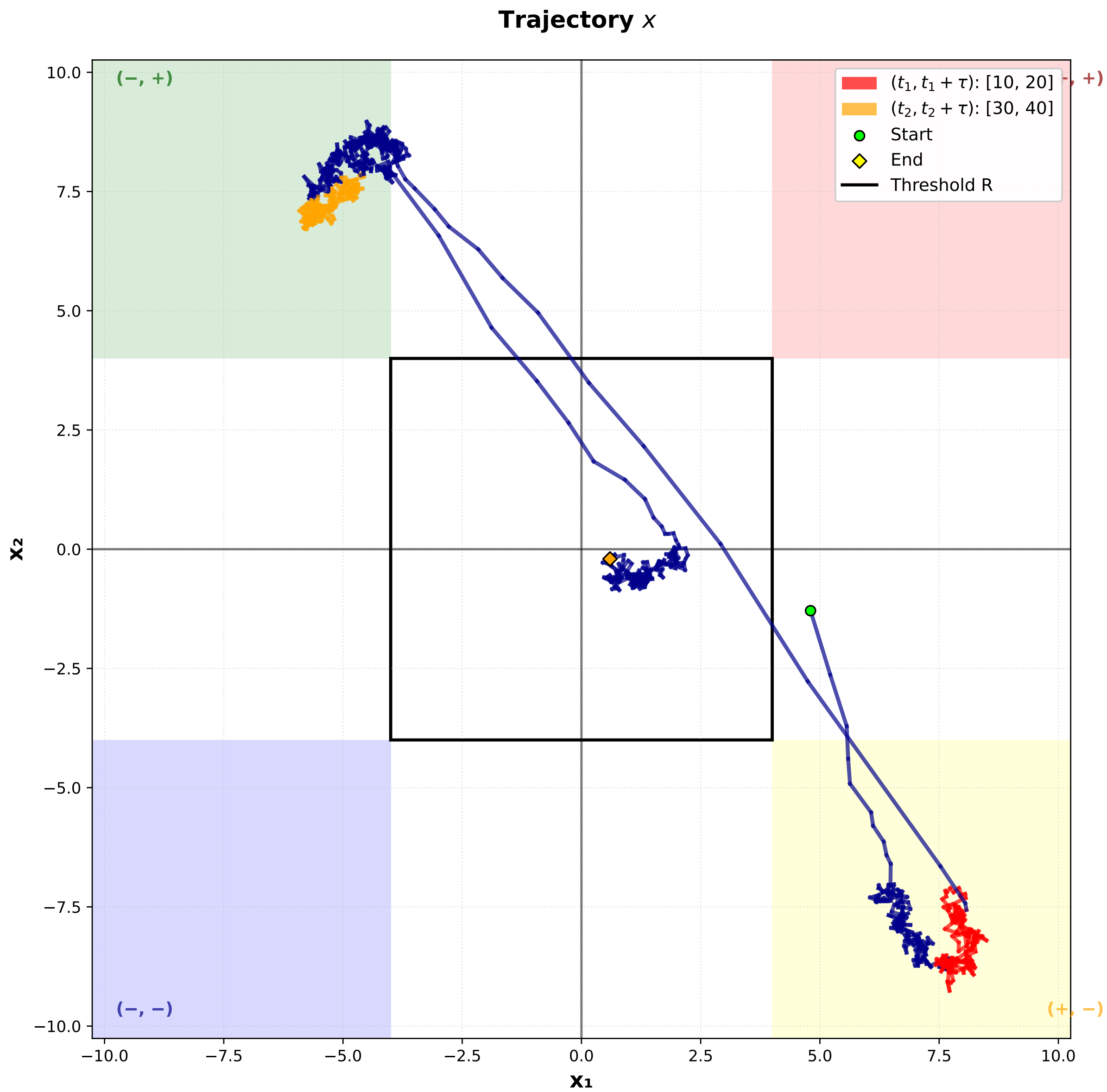


Figure 8.1: Trajectory of x in the 2-dimensional phase plane. The trajectory is color-coded by saturation regime: the red segment corresponds to the interval $[t_1, t_1 + \tau]$ during which $S(x(t)) = v_1 = (1, 0)^\top$, and the yellow segment corresponds to $[t_2, t_2 + \tau]$ during which $S(x(t)) = v_2 = (0, 1)^\top$. The gray portions are transient phases between the two saturated regimes. The two highlighted segments provide the independent excitation directions required for the algebraic estimator E_2 to recover the full matrix W.

9

INFINITE-DIMENSIONAL SYSTEM

In the previous chapters, we developed the estimation framework for saturated systems in finite dimension. Given access to the saturated output $y = WS(V)$, a sequence of open-loop controls steers the saturation to canonical elements of $\mathbb{R}^n$ from which the unknown output matrix W is recovered exactly. In this chapter, we extend this framework to the infinite-dimensional setting $\mathcal{H} = L^2(\Omega)$, motivated by applications to neural field models of the form $\partial_t V = -V + A \star S(V) + u$. The main challenge is that an infinite-dimensional operator seems difficult to be recovered exactly from finitely many experiments; instead, we seek an approximation to any desired accuracy $\varepsilon > 0$. We show that this is achievable: by partitioning Ω into dyadic cells and steering $S(V(\cdot, t))$ to each indicator function $\mathbf{1}_{Q_k}$ in turn, one recovers the action of W on theses indicator functions, yielding a finite-rank approximation $\hat{W}$ whose error depends on the mesh width of the dyadic decomposition, assuming some Sobolev regularity of W. The analysis proceeds in three stages: a scalar dissipative case where the pointwise argument of the finite-dimensional theory carries over directly; an extension to bounded spatial coupling, known-linear-part extension where the sign structure of the resolvent $(-\mathcal{A})^{-1}B$ replaces the small-coupling condition. Finally, we discuss some perspectives on the extension to more general structure with diffusions operatoros.

9.1 INTRODUCTION

This chapter extends the estimation results of Chapter 8 to an infinite-dimensional setting, completing the work on estimation of saturated networks. The extension is motivated by the desire to apply the developed techniques to neural fields as described in the introduction, which are naturally modelled as infinite-dimensional systems.

The central difficulty in passing to infinite dimension is that an infinite-dimensional operator W seems difficult to be recovered exactly from a finite number of experiments. Any estimation procedure must therefore involve an approximation, and the question becomes: can we design a control strategy that recovers W up to any prescribed accuracy $\varepsilon > 0$? We show that this is indeed possible, and that the approximation error can be made explicit in terms of the regularity of W and the resolution of a dyadic partition of Ω.

The chapter is organized as follows. Section 9.2 provides an informal overview of the approach and the main ideas. Section 9.3 states the main result (Theorem 9.7) for the scalar dissipative case $\mathcal{A} = -\lambda I$. Sections 9.4 and 9.5 collect the well-posedness and technical lemmas. Section 9.7 extends the result to bounded operators satisfying a small-coupling condition $\kappa < 1$.

9.2 HEURISTICS

The main motivation is to estimate parameters in neural fields. Consider a neural field whose state takes values in a Hilbert space $\mathcal{H}$ whose elements are real valued functions, governed by a system of the form

$$\begin{cases} \dot{V} = -V + A \star S(V) + u, \\ y(x,t) = WS(V(x,t)), \end{cases} \tag{9.1}$$

where $S : \mathbb{R} \to \mathbb{R}$ is the saturation function as defined in Chapter 8. Since W is infinite-dimensional, exact recovery from finitely many experiments seems difficult for a general operator. The question is whether the accuracy can be made arbitrarily small. The answer is yes: by steering the saturation output to successive indicator functions over disjoint time intervals, one recovers the action of W on a dyadic basis exactly, yielding an approximation to any prescribed precision.

To illustrate this concretely, suppose the output takes the form $y(x,t) = \int_\Omega K(x,s)S(V(s,t))\,ds$ for some kernel K. If we can design a control such that $S(V(\cdot,t))$ equals the indicator function of a cell Q_k in a partition of Ω, then the output directly reveals the average of K over Q_k:

$$\int_\Omega K(x,s)\underbrace{S(V(s,t))}_{=\mathbf{1}_{Q_k}(s)}\,ds = \int_{Q_k} K(x,s)\,ds. \tag{9.2}$$

Should the size of the cells be sufficiently small, based on the output y, one could then construct an estimate $\hat{K}$ which equals to the average of K over each cell.

To turn this idea into a rigorous approximation scheme, we exploit the fact that step functions on dyadic partitions are dense in $L^2(\Omega)$. We partition Ω using the ambient dyadic grid of $\mathbb{R}^d$ at a given resolution level j: the dyadic cubes of side $h = 2^{-j}$, intersected with Ω, cover the domain up to a null set. The normalized indicator functions of these cells form an orthonormal system, and the approximation error is controlled by the mesh width h.

The procedure then consists of the following steps. Fix a dyadic level $j \in \mathbb{N}$ and let $\{Q_1, \dots, Q_N\}$ denote the dyadic partition of Ω at level j (see Lemma 9.5). First, we prove that for any prescribed time interval, $S(V)$ can be made equal to the indicator function of any cell Q_k. This is achieved by using a constant control that is positive and large enough inside Q_k, and negative and large enough outside. Then, we repeat this process to ensure that $S(V)$ equals $\mathbf{1}_{Q_k}$ for $k = 1, \dots, N$ on disjoint time intervals. On each of these time intervals, the output equals $y(t) = W\mathbf{1}_{Q_k}$, from which the action of W on the basis element $\tilde{e}_k = |Q_k|^{-1/2}\mathbf{1}_{Q_k}$ is recovered *exactly*.

After all N experiments, one obtains the finite-rank approximation $\hat{W}_j = W \circ P_j$, where P_j is the orthogonal projection onto the span of the indicator functions. The error $\|W - \hat{W}_j\|$ is purely a projection error, controlled by the mesh width $h = 2^{-j}$. Under regularity assumptions on W, this error is of order $O(h)$.

The remaining question is to quantify the approximation error as a function of the chosen dyadic level j. The projection P_j replaces W by its cell-by-cell averages; the error $W - \hat{W}_j$ measures how much W varies *within* each cell. To control this, we place ourselves in the Sobolev space $H^1(\Omega)$ and exploit the Poincaré–Wirtinger inequality: on any domain of diameter h, the difference between a function and its mean is bounded (in L^2) by the width of the dyadic cells (2^{-j}) times its gradient. Applied cell by cell to the dyadic partition at level j, this yields an approximation error of order $O(h) = O(2^{-j})$, uniform in time and explicit in terms of the H^1-regularity of W.

9.3 MAIN RESULT

We set ourselves in the Hilbert space $\mathcal{H} = L^2(\Omega)$, with $\Omega \subset \mathbb{R}^d$ a bounded Lipschitz domain. Consider the following dynamical system:

$$\begin{cases} \partial_t V(x,t) = -\lambda V(x,t) + F(V,t)(x) + u(x,t), \quad x \in \Omega, \\ y = \mathcal{W}[S(V(\cdot,t))] \end{cases} \tag{9.3}$$

where $V(\cdot,t) \in L^\infty(\Omega)$ is the (unknown) state, $u(\cdot,t) \in L^\infty(\Omega)$ is the spatially distributed control, F is a nonlinear bounded perturbation (see Assumption 9.2 below), and S is the saturation function applied pointwise $S(V)(x) := S(V(x))$ (Assumption 9.1). The output operator $\mathcal{W} = W : L^2(\Omega) \to \mathcal{G}$ is an unknown bounded linear operator into a separable Hilbert space $\mathcal{G}$, yielding $y(\cdot,t) \in \mathcal{G}$.

Assumption 9.1. *The saturation function $S : \mathbb{R} \to \mathbb{R}$ is known, globally Lipschitz-continuous, and piecewise C^1. There exist known constants $\underline{R} < \bar{R}$ such that*

$$x \leq \underline{R} \Rightarrow S(x) = 0, \qquad x \geq \bar{R} \Rightarrow S(x) = 1,$$

and S is strictly increasing on $[\underline{R}, \bar{R}]$. It is applied pointwise: $S(V)(x) := S(V(x))$. We denote $R := \max\{|\underline{R}|, |\bar{R}|\}$.

Assumption 9.2. *The map $F : L^2(\Omega) \times \mathbb{R}_+ \to L^\infty(\Omega)$ satisfies the uniform bound*

$$\|F(V,t)\|_{L^\infty(\Omega)} \leq b_F,$$

for a known constant $b_F > 0$, for all $V \in L^2(\Omega)$ and all $t \geq 0$.

Assumption 9.3. *The control $u : \Omega \times \mathbb{R}_+ \to \mathbb{R}$ is bounded and piecewise constant in time: $u \in L^\infty([0,T], L^\infty(\Omega))$ for every $T > 0$.*

The meaning and assumptions on each term are summarized in Table 9.1.

Remark 9.4. *The system (9.3) captures in particular the neural field model $\partial_t V = -V + A \star S(V) + u$ by setting $\lambda = 1$ and $F(V)(x) = (A \star S(V))(x)$ with $A \in L^1(\Omega)$. Since $\|S(V)\|_{L^\infty} \leq 1$, Young's convolution inequality gives $\|F(V)\|_{L^\infty} \leq \|A\|_{L^1}\|S(V)\|_{L^\infty} \leq \|A\|_{L^1}$, so the boundedness assumption is automatically satisfied with $b_F = \|A\|_{L^1}$. More generally, the dissipative term $-\lambda V$ may be replaced by a linear operator $\mathcal{A}$ generating an exponentially stable C_0-semigroup on $L^2(\Omega)$, provided its semigroup satisfies a pointwise decay estimate of the form $|(e^{t\mathcal{A}}\phi)(x)| \leq Me^{-\lambda t}|\phi(x)|$ for a.e. $x \in \Omega$.*

Term	**Unknown**	**To Retrieve?**	**Meaning**	**Assumptions**
V	✓	Yes	State field	$V(\cdot,t) \in L^\infty(\Omega)$. Unknown; evolves according to the system dynamics.
u			Control input	Spatially distributed, bounded and piecewise constant in time: $u \in L^\infty([0,T], L^\infty(\Omega))$ (Assumption 9.3).
y			Measured output	Noise free.
λ			Damping rate	Known constant $\lambda > 0$. Ensures that the linear part $-\lambda V$ is dissipative (Hurwitz).
F	✓		Nonlinear bounded perturbation	$F : L^2(\Omega) \times \mathbb{R}_+ \to L^\infty(\Omega)$, uniformly bounded: $\sup_t \|F(V,t)\|_{L^\infty} \leq b_F$ for a known $b_F > 0$ (Assumption 9.2).
W	✓	Yes	Unknown output operator	Hilbert–Schmidt operator $W : L^2(\Omega) \to \mathcal{G}$, with known bound $\|W\|_{\mathrm{HS}} \leq C_W$; to be estimated.
S			Saturation function	Assumption 9.1: known, globally Lipschitz-continuous, piecewise C^1, with known saturation levels $\underline{R} < \bar{R}$ such that $x \leq \underline{R} \Rightarrow S(x) = 0$ and $x \geq \bar{R} \Rightarrow S(x) = 1$, strictly increasing on $[\underline{R}, \bar{R}]$. Applied pointwise: $S(V)(x) = S(V(x))$. Denote $R := \max\{\lvert\underline{R}\rvert, \bar{R}\}$.

Table 9.1: Summary of variables and assumptions for the infinite-dimensional system.

We decompose Ω into dyadic cells and use the saturation function to generate, on disjoint time intervals, the indicator functions of these cells. We fix a dyadic level $j \in \mathbb{N}$ and partition Ω into dyadic cells, as described by the following lemma. This is a standard construction which can be found in classic textbook of measure theory and functional analysis or finite element literature [37].

Lemma 9.5 (Dyadic partition of $L^2(\Omega)$)**.** *Let $\Omega \subset \mathbb{R}^d$ be a bounded open set. For each level $j \in \mathbb{N}$, consider the dyadic cubes of $\mathbb{R}^d$ at level j:*

$$Q_{j,\mathbf{m}} = \prod_{i=1}^{d} \big[m_i\, 2^{-j},\ (m_i{+}1)\, 2^{-j}\big), \quad \mathbf{m} \in \mathbb{Z}^d,$$

and define the dyadic partition of Ω at level j *by*

$$\mathcal{D}_j(\Omega) = \big\{ Q_{j,\mathbf{m}} \cap \Omega \ :\ \mathbf{m} \in \mathbb{Z}^d,\ |Q_{j,\mathbf{m}} \cap \Omega| > 0 \big\} =: \{Q_1, \dots, Q_{N_j}\}.$$

The cells Q_k are pairwise disjoint, cover Ω up to a null set, and have diameter at most $h\sqrt{d}$ where $h := 2^{-j}$ is the mesh width. *The number of cells satisfies $N_j \le |\Omega|\, h^{-d} + C_{\partial\Omega}\, h^{-d+1}$, where the second term accounts for cells intersecting $\partial\Omega$.*

The normalized indicator functions

$$\tilde{e}_k := |Q_k|^{-1/2}\, \mathbf{1}_{Q_k}, \quad k = 1, \dots, N_j, \tag{9.4}$$

form an orthonormal system in $L^2(\Omega)$, and the orthogonal projection

$$P_j f := \sum_{k=1}^{N_j} \langle f, \tilde{e}_k \rangle\, \tilde{e}_k = \sum_{k=1}^{N_j} \frac{1}{|Q_k|} \left(\int_{Q_k} f\, dx \right) \mathbf{1}_{Q_k} \tag{9.5}$$

satisfies $P_j f \to f$ in $L^2(\Omega)$ as $j \to \infty$ for every $f \in L^2(\Omega)$.

Moreover, for $f \in H^1(\Omega)$, using the Poincaré–Wirtinger inequality and following classical finite element approaches such as in [37, Theorem 4.4.20], one has the bound:

$$\|f - P_j f\|_{L^2(\Omega)} \le C_P\, h\, \|\nabla f\|_{L^2(\Omega)}, \tag{9.6}$$

where $C_P > 0$ depends only on the dimension d. We recall that $h = 2^{-j}$ is the mesh width of the dyadic partition.

Proof. The dyadic cubes $\{Q_{j,\mathbf{m}}\}_{\mathbf{m} \in \mathbb{Z}^d}$ partition $\mathbb{R}^d$ (up to the boundaries, which form a null set). Intersecting with Ω yields pairwise disjoint sets that tile Ω up to a null set, since $\partial Q_{j,\mathbf{m}}$ has measure zero for each $\mathbf{m}$, and only finitely many cubes intersect the bounded set Ω.

Orthonormality is immediate: $\langle \tilde{e}_k, \tilde{e}_l \rangle = |Q_k|^{-1/2} |Q_l|^{-1/2} \int_{Q_k \cap Q_l} dx = \delta_{kl}$ since the cells are disjoint.

For the strong convergence $P_j f \to f$: at a fixed level j, note that $P_j f$ is the approximation of f using step functions of mesh width $h = 2^{-j}$. As j increases, $h \to 0$ and the step functions at level j become increasingly fine. Indeed, every $f \in L^2(\Omega)$ can be approximated arbitrarily closely by a step function constant on dyadic cells of sufficiently small side. This is the well known and standard results on the density of step functions in L^p spaces. This result can be seen in standard measure theory books such as [81].

For the quantitative bound (9.6), on each cell Q_k, the Poincaré–Wirtinger inequality gives

$$\int_{Q_k} |f - \bar{f}_k|^2 \, dx \le C_P^2 \, (\operatorname{diam} Q_k)^2 \int_{Q_k} |\nabla f|^2 \, dx,$$

where $\bar{f}_k = |Q_k|^{-1} \int_{Q_k} f \, dx$ is the average of f on Q_k. Since $P_j f|_{Q_k} = \bar{f}_k$ and $\operatorname{diam} Q_k \le h\sqrt{d}$, summing over all cells:

$$\|f - P_j f\|_{L^2}^2 = \sum_{k=1}^{N_j} \int_{Q_k} |f - \bar{f}_k|^2 \, dx \le C_P^2 \, d \, h^2 \sum_{k=1}^{N_j} \int_{Q_k} |\nabla f|^2 \, dx = C_P^2 \, d \, h^2 \, \|\nabla f\|_{L^2}^2.$$

Absorbing the dimensional constant into C_P gives (9.6). □

Remark 9.6. *The dyadic partition $\mathcal{D}_j(\Omega)$ at level j is a particular instance of a* non-degenerate family of subdivisions $\{T_h\}_{0<h\le 1}$ *of Ω in the sense of [37, Definition 4.4.13], with mesh parameter $h = 2^{-j}$. The interior cells are identical cubes of side h, the most regular possible mesh. The boundary cells $Q_{j,\mathbf{m}} \cap \Omega$ are cut by $\partial\Omega$ and may be irregular, but they remain connected Lipschitz domains. Indeed, the intersection of a convex set (the cube $Q_{j,\mathbf{m}}$) with a Lipschitz domain (Ω) is itself Lipschitz. The Poincaré–Wirtinger inequality therefore holds on each cell, including the boundary cells, and the bound (9.6) follows from [37, Theorem 4.4.20] applied to this family.*

We henceforth fix a dyadic level j and write $(Q_k, \tilde{e}_k)_{k=1}^{N_j}$ for the dyadic partition of Ω at level j as in Lemma 9.5, with mesh width $h = 2^{-j}$ with the corresponding step functions $\tilde{e}_k = |Q_k|^{-1/2} \mathbf{1}_{Q_k}$. The key simplification compared to using an abstract Hilbert basis (e_i) is that each basis element $\tilde{e}_k$ is itself a (scaled) indicator function of a dyadic cell, so a *single* indicator-realization experiment directly produces the corresponding projection coefficient of $\mathcal{W}$, with no density approximation step.

Theorem 9.7. *Under the assumptions of Table 9.1, fix a dyadic level $j \in \mathbb{N}$ and let $(Q_k, \tilde{e}_k)_{k=1}^{N_j}$ be the dyadic partition of Ω at level j with mesh width $h = 2^{-j}$ (Lemma 9.5). For any $\rho > 0$, $\tau > 0$, and any time instants $0 < t_1 < t_2 < \cdots < t_{N_j}$ satisfying $t_1 \ge \tau$ and $t_{k+1} > t_k + \tau$ for all k, there exists a bounded piecewise constant control $u^\star \in L^\infty([0, T], L^\infty(\Omega))$ with $T = t_{N_j} + \tau$, such that the estimator*

$$\hat{y}_k := \frac{1}{\tau |Q_k|^{1/2}} \int_{t_k}^{t_k+\tau} y(s) \, ds, \quad k = 1, \ldots, N_j, \tag{9.7}$$

satisfies the following for any solution of (9.3) initialized with $\|V(\cdot, 0)\|_{L^\infty(\Omega)} \le \rho$: for all $k = 1, \ldots, N_j$,

$$\hat{y}_k = W \frac{\mathbf{1}_{Q_k}}{|Q_k|^{1/2}} = W\tilde{e}_k. \tag{9.8}$$

The finite-rank operator $\hat{W}_j : L^2(\Omega) \to \mathcal{G}$ defined by

$$\hat{W}_j(\phi) := \sum_{k=1}^{N_j} \langle \phi, \tilde{e}_k \rangle \, \hat{y}_k \tag{9.9}$$

satisfies

$$\|W - \hat{W}_j\|_{\mathrm{op}} \xrightarrow{j\to\infty} 0. \tag{9.10}$$

In particular, the images $W\tilde{e}_k$ are recovered exactly *from the experiment, with no approximation error. The error is purely a* projection error *controlled by the mesh width. By Lemma 9.5,*

$$\|W - \hat{W}_j\|_{\mathrm{op}} \le C_P\, h\, \|W^*\|_{\mathcal{G}\to H^1}, \tag{9.11}$$

provided that $W^* : \mathcal{G} \to H^1(\Omega)$ is bounded (see Subsection 9.3.1).

As $j \to \infty$ (i.e. $h \to 0$), $\hat{W}_j \to W$ in operator norm for any compact W, with no regularity assumption.

Remark 9.8. *When W is additionally Hilbert–Schmidt, the Pythagorean theorem gives*

$$\|W\|_{\mathrm{HS}}^2 = \sum_{k=1}^{N_j} \|W\tilde{e}_k\|_{\mathcal{G}}^2 + \|W|_{E_j^\perp}\|_{\mathrm{HS}}^2, \tag{9.12}$$

where $E_j^\perp = (\mathrm{span}\{\tilde{e}_1, \dots, \tilde{e}_{N_j}\})^\perp$. Since $W\tilde{e}_k = \hat{y}_k$ is recovered exactly, the residual $\|W|_{E_j^\perp}\|_{\mathrm{HS}}^2 = \|W\|_{\mathrm{HS}}^2 - \sum_{k=1}^{N_j} \|\hat{y}_k\|_{\mathcal{G}}^2$ is determined by the data, and $\|W - \hat{W}_j\|_{\mathrm{op}} \le \|W|_{E_j^\perp}\|_{\mathrm{HS}}$.

Remark 9.9. *The estimator (9.7) is the time-average of the output over an interval during which $S(V(\cdot,t)) = \mathbf{1}_{Q_k}$ exactly. Each experiment directly yields the action of $\mathcal{W}$ on these elements. This is the infinite-dimensional analog of the algebraic estimator E_2 in Chapter 8, where distinct saturation values v_i were used to generate linearly independent excitation directions.*

Remark 9.10. *The total number of experiments equals the number of dyadic cells at level j: $N_j \lesssim |\Omega|\, h^{-d}$. The experiment duration is $T = N_j(\tau + T_0)$ where T_0 is the settling time of the indicator realization (Lemma 9.14 below). The growth of N_j with the dimension d can be managed as most systems of interest are low-dimensional (e.g., $d = 1$ or $d = 2$ for neural fields).*

9.3.1 *Regularity assumptions and error bounds*

The error bound (9.11) admits an $O(h)$ dependence on the mesh width h under regularity assumptions on W. This H^1 regularity is necessary. Indeed, the naïve factorization $\|W(I - P_j)\| \le \|W\|\,\|I - P_j\|$ is useless since $I - P_j$ is an orthogonal projection onto an infinite-dimensional subspace of $L^2(\Omega)$; one has $\|I - P_j\|_{L^2\to L^2} = 1$ for every j, regardless of the mesh width. However, as a map from the *smoother* space $H^1(\Omega)$ to $L^2(\Omega)$, the projection error does decay:

$$\|I - P_j\|_{H^1(\Omega)\to L^2(\Omega)} \le C_P\, h, \tag{9.13}$$

which is exactly the content of the Poincaré–Wirtinger inequality applied cell by cell. The quantitative error bound then follows by factoring through the adjoint:

$$\|W - \hat{W}_j\|_{\mathrm{op}} = \|W(I - P_j)\|_{L^2\to\mathcal{G}} = \|(I - P_j)W^*\|_{\mathcal{G}\to L^2} \le \|I - P_j\|_{H^1\to L^2}\, \|W^*\|_{\mathcal{G}\to H^1} \le C_P\, h\, \|W^*\|_{\mathcal{G}\to H^1}. \tag{9.14}$$

The first equality uses the adjoint identity $\|T\|_{\mathrm{op}} = \|T^*\|_{\mathrm{op}}$ and the self-adjointness of $I - P_j$. The inequality is a composition bound, valid precisely when W^* maps $\mathcal{G}$ into $H^1(\Omega)$ boundedly:

$$\|W^*\|_{\mathcal{G}\to H^1} := \sup_{\|g\|_{\mathcal{G}}=1} \|W^* g\|_{H^1(\Omega)} < \infty, \tag{9.15}$$

where $\|f\|_{H^1}^2 = \|f\|_{L^2}^2 + \|\nabla f\|_{L^2}^2$ is the standard Sobolev norm. This is the *minimal* assumption for the $O(h)$ rate.

In the scalar case $\mathcal{G} = \mathbb{R}$, condition (9.15) reduces to $w \in H^1(\Omega)$ with $\|W^*\|_{\mathbb{R}\to H^1} = \|w\|_{H^1}$, where $w \in L^2(\Omega)$ is the Riesz representative of W defined by $Wf = \langle w, f\rangle_{L^2}$.

Condition (9.15) admits two equivalent formulations:

(i) *Kernel form.* If W is an integral operator with kernel $K(x,y)$, the condition reads: $K(x,\cdot) \in H^1(\Omega)$ for a.e. $x \in \Omega'$. Differentiating under the integral gives $\nabla_y W^* g(y) = \int_{\Omega'} g(x)\, \nabla_y K(x,y)\, dx$, so by Cauchy–Schwarz (with $\mathcal{G} = L^2(\Omega')$):

$$\|W^*\|_{\mathcal{G}\to H^1}^2 \le \int_{\Omega'}\int_{\Omega} \left(|K(x,y)|^2 + |\nabla_y K(x,y)|^2\right) dy\, dx. \tag{9.16}$$

Hence, the natural sufficient condition is $\nabla_y K \in L^2(\Omega' \times \Omega)$, which also makes W^* Hilbert–Schmidt as a map $L^2(\Omega') \to H^1(\Omega)$.

(ii) *Singular value form.* If $(\sigma_k, \phi_k, \psi_k)$ is the singular system of W (i.e. $W\phi_k = \sigma_k \psi_k$, $W^*\psi_k = \sigma_k\phi_k$), then $W^* g = \sum_k \sigma_k \langle g, \psi_k\rangle \phi_k$, so $\nabla W^* g = \sum_k \sigma_k \langle g, \psi_k\rangle \nabla\phi_k$. The natural sufficient condition is

$$\sum_k \sigma_k^2\, \|\nabla\phi_k\|_{L^2}^2 < \infty, \tag{9.17}$$

which gives $\|W^*\|_{\mathcal{G}\to H^1}^2 \le \sum_k \sigma_k^2 (1 + \|\nabla\phi_k\|_{L^2}^2) < \infty$. Geometrically, this requires the left singular functions ϕ_k to remain in $H^1(\Omega)$ with $\|\nabla\phi_k\|_{L^2}$ not growing faster than σ_k^{-1}.

9.4 EXISTENCE AND BOUNDEDNESS OF SOLUTIONS

Before proceeding to the proof of the main theorem, we establish the well-posedness of the system. The scalar case (9.3) is a particular instance of the general semigroup formulation

$$\partial_t V = \mathcal{A}V + F(V,t) + u, \tag{9.18}$$

with $\mathcal{A} = -\lambda I$, which generates the C_0-semigroup $e^{t\mathcal{A}} = e^{-\lambda t} I$ on $L^2(\Omega)$. We therefore state the well-posedness result directly for a general semigroup generator $\mathcal{A}$. The scalar case follows as a specific case. Note that for simplicity, we denote the semigroup as $e^{t\mathcal{A}}$.

Assumption 9.11. *The operator $\mathcal{A} : D(\mathcal{A}) \subset L^2(\Omega) \to L^2(\Omega)$ generates an exponentially stable C_0-semigroup $(e^{t\mathcal{A}})_{t\ge 0}$ on $L^2(\Omega)$: there exist constants $M \ge 1$ and $\omega > 0$ such that*

$$\|e^{t\mathcal{A}}\|_{L^2\to L^2} \le M e^{-\omega t}, \quad \forall t \ge 0. \tag{9.19}$$

Moreover, the semigroup extends to $L^\infty(\Omega)$ *with*

$$\|e^{t\mathcal{A}}\|_{L^\infty\to L^\infty} \le M_\infty\, e^{-\omega_\infty t}, \quad \forall t \ge 0, \tag{9.20}$$

for constants $M_\infty \ge 1$ *and* $\omega_\infty > 0$.

This is a specific setting for the standard results on existence of solutions for evolution equations of type (9.18). The following proposition and its proof follow [134, Chapter 6/Theorem4.3]

Proposition 9.12 (Well-posedness). *Under Assumption 9.11, let* $F : L^2(\Omega)\times\mathbb{R}_+ \to L^2(\Omega)$ *satisfy* $\|F(V,t)\|_{L^\infty} \le b_F$ *for all* (V,t), *and let* $u \in L^\infty([0,T], L^\infty(\Omega))$. *For any* $V_0 \in L^2(\Omega)\cap L^\infty(\Omega)$, *equation (9.18) admits a unique mild solution*

$$V(\cdot,t) = e^{t\mathcal{A}}V_0 + \int_0^t e^{(t-s)\mathcal{A}}\big[F(V,s) + u(\cdot,s)\big]\,ds \tag{9.21}$$

in $C([0,T], L^2(\Omega))$ *for any* $T > 0$. *Moreover,* $V \in L^\infty([0,T], L^\infty(\Omega))$ *with the uniform bound*

$$\|V(\cdot,t)\|_{L^\infty} \le M_\infty\,\|V_0\|_{L^\infty}\, e^{-\omega_\infty t} + \frac{M_\infty(b_F + \|u\|_{L^\infty})}{\omega_\infty}, \quad \forall t \ge 0. \tag{9.22}$$

In particular, the solution is global in time. If additionally F is Lipschitz in V (e.g., $F(V) = A \star S(V)$ *with S globally Lipschitz), then the mild solution is unique without further conditions.*

Proof. The existence and uniqueness of mild solutions in $C([0,T], L^2(\Omega))$ follows from standard semigroup theory [134, Chapter 6/Theorem4.3]. The term $G(V,t) := F(V,t) + u(\cdot,t)$ satisfies $\|G(V,t)\|_{L^\infty} \le b_F + \|u\|_{L^\infty}$. If F is Lipschitz in V, the Banach fixed-point theorem yields existence and uniqueness directly. If F is merely bounded, a Schauder-type fixed-point argument applies to ensure existence as it is L^1_{loc} in time.

The L^∞-estimate follows from the L^∞-boundedness of the semigroup:

$$\|V(\cdot,t)\|_{L^\infty} \le M_\infty e^{-\omega_\infty t}\|V_0\|_{L^\infty} + \int_0^t M_\infty e^{-\omega_\infty(t-s)}(b_F + \|u\|_{L^\infty})\,ds,$$

and the integral evaluates to $M_\infty(b_F + \|u\|_{L^\infty})/\omega_\infty$. The bound is uniform in time, so the solution exists globally. □

The following regularity result is classical; it is a direct consequence of the theory of uniformly continuous semigroups and Bochner-space calculus (see [134, Section 4.2, Corollary 4.2.5] and [71, Section II.6]).

Proposition 9.13 (Regularity of mild solutions for bounded generators). *Under the above assumptions and Assumptions 9.2 and 9.3. Then the mild solution (9.21) is a* strong solution, *i.e. V is absolutely continuous from* $[0,T]$ *to* $L^\infty(\Omega)$, $V(t) \in D(\mathcal{A}) = L^\infty(\Omega)$ *for all* $t \ge 0$, *and*

$$\frac{d}{dt}V(\cdot,t) = \mathcal{A}\,V(\cdot,t) + F(V(\cdot,t),t) + u(\cdot,t) \quad \text{in } L^\infty(\Omega), \quad \text{for a.e. } t \in [0,T]. \tag{9.23}$$

In particular, for a.e. $(x,t) \in \Omega\times[0,T]$, *the pointwise identity*

$$\partial_t V(x,t) = (\mathcal{A}\,V(\cdot,t))(x) + F(V(\cdot,t),t)(x) + u(x,t) \tag{9.24}$$

holds. If additionally u and $t \mapsto F(V(\cdot,t),t)$ *are continuous in time, then* $V \in C^1([0,T], L^\infty(\Omega))$ *and the identity holds for all t.*

9.5 TECHNICAL LEMMAS FOR THE PROOF OF THE MAIN THEOREM

In this section, we present the technical results that underpin the proof of Theorem 9.7. The main building block is the ability to steer $S(V(\cdot,t))$ to any prescribed indicator function of a measurable subset of Ω, uniformly over all initial conditions in a bounded set. The well-posedness of the system established in Section 9.4 guarantees that the solutions are globally defined and bounded in $L^\infty(\Omega)$. The first lemma establishes the pointwise open-loop controllability of the saturation output to any indicator function. The argument is fundamentally pointwise: the distributed control acts independently at each spatial location, and the bounded perturbation can be dominated by choosing the control amplitude large enough.

Lemma 9.14 (Indicator function realization)**.** *Let $\rho > 0$ and $T_0 > 0$. Let $Q \subset \Omega$ be a measurable set with $|Q| > 0$. There exists $\mu^\star > 0$ such that for any $\mu \geq \mu^\star$, the constant-in-time control*

$$u_\mu(x) = \mu\,\mathbf{1}_Q(x) - \mu\,\mathbf{1}_{\Omega\setminus Q}(x) \tag{9.25}$$

yields, for any solution of (9.3) initialized with $\|V(\cdot,0)\|_{L^\infty(\Omega)} \leq \rho$:

$$S(V(x,t)) = \mathbf{1}_Q(x), \quad \text{for a.e. } x \in \Omega, \quad \forall t \geq T_0. \tag{9.26}$$

Moreover, the state satisfies the uniform-in-time bound

$$\|V(\cdot,t)\|_{L^\infty} \leq \rho + \frac{b_F + \mu}{\lambda}, \quad \forall t \geq 0. \tag{9.27}$$

Proof. The argument proceeds pointwise by exploiting the scalar dissipative structure. For a.e. $x \in Q$, the dynamics satisfy

$$\partial_t V(x,t) = -\lambda V(x,t) + F(V,t)(x) + \mu \geq -\lambda V(x,t) - b_F + \mu. \tag{9.28}$$

By the comparison principle for the scalar ODE $\dot{w} = -\lambda w + (\mu - b_F)$, using $V(x,0) \geq -\rho$:

$$V(x,t) \geq -\rho\, e^{-\lambda t} + \frac{\mu - b_F}{\lambda}(1 - e^{-\lambda t}). \tag{9.29}$$

For $t \geq T_0$, the right-hand side is at least $-\rho\, e^{-\lambda T_0} + \frac{\mu - b_F}{\lambda}(1 - e^{-\lambda T_0})$. Setting

$$\mu_1^\star := b_F + \frac{\lambda(\bar{R} + \rho)}{1 - e^{-\lambda T_0}} \tag{9.30}$$

ensures that $V(x,t) \geq \bar{R}$ for all $t \geq T_0$. Moreover, since $V(x,t) \to (\mu - b_F)/\lambda > \bar{R}$ as $t \to \infty$, the dissipative term ensures that $V(x,t)$ remains above $\bar{R}$ for all $t \geq T_0$, and hence $S(V(x,t)) = 1$ by the saturation assumption.

Similarly, for a.e. $x \in \Omega \setminus Q$:

$$\partial_t V(x,t) = -\lambda V(x,t) + F(V,t)(x) - \mu \leq -\lambda V(x,t) + b_F - \mu. \tag{9.31}$$

By comparison: $V(x,t) \leq \rho\, e^{-\lambda t} + \frac{b_F - \mu}{\lambda}(1 - e^{-\lambda t})$. Setting $\mu_2^\star := b_F + \frac{\lambda(|\underline{R}|+\rho)}{1-e^{-\lambda T_0}}$ ensures $V(x,t) \leq \underline{R}$ for all $t \geq T_0$, and by the same dissipative argument the bound is maintained for all subsequent times, giving $S(V(x,t)) = 0$.

Setting $\mu^\star := \max\{\mu_1^\star, \mu_2^\star\} = b_F + \frac{\lambda(R+\rho)}{1-e^{-\lambda T_0}}$ with $R = \max\{|\underline{R}|, \bar{R}\}$ establishes (9.26). For the uniform bound (9.27), the comparison principle gives for all $t \geq 0$:

$$|V(x,t)| \leq \rho\, e^{-\lambda t} + \frac{b_F + \mu}{\lambda}(1 - e^{-\lambda t}) \leq \rho + \frac{b_F + \mu}{\lambda}, \tag{9.32}$$

which is uniform in time thanks to the dissipative term $-\lambda V$. □

Remark 9.15. *The control in Lemma 9.14 is piecewise constant, $+\mu$ on Q and $-\mu$ on $\Omega \setminus Q$. In practice one may replace it by a smooth approximation via a mollifier. The realization is then no longer exact, but one still obtains $S(V(x,t)) \approx \mathbf{1}_Q(x)$ for $t \geq T_0$, with an error controlled by the mollification width. This situation is similar to the unbounded-operator case discussed in Section 9.9, where a boundary layer of width δ_0 prevents exact realization near ∂Q.*

Remark 9.16. *In the lemma above we exploited the fact that the control acts independently at each spatial location and that the operator $-\lambda I$ acts component wise. Unlike the cone-based geometry of Section 8.7, where steering into different cones required coordinating all components simultaneously, the control here acts independently at each spatial location. This effectively reduces the problem to a family of scalar dissipative sub-problems indexed by $x \in \Omega$.*

The following lemma extends the indicator realization to a sequence of prescribed indicator functions realized on disjoint time intervals.

Lemma 9.17 (Sequential indicator function realization)**.** *Let $\rho > 0$, $\tau > 0$, and $N \in \mathbb{N}$. Let $(Q_k)_{1 \leq k \leq N}$ be measurable subsets of Ω with $|Q_k| > 0$ for all k. Let $t_1 < t_2 < \cdots < t_N$ be time instants with $t_{k+1} > t_k + \tau$ for all k and $t_1 > \tau$. There exists a piecewise constant control $u^\star \in L^\infty([0,T], L^\infty(\Omega))$ with $T = t_N + \tau$ such that for any solution of (9.3) initialized with $\|V(\cdot,0)\|_{L^\infty} \leq \rho$:*

$$S(V(x,t)) = \mathbf{1}_{Q_k}(x), \quad \text{a.e. } x \in \Omega, \quad \forall t \in [t_k, t_k + \tau], \quad k = 1, \ldots, N. \tag{9.33}$$

Proof. The proof proceeds by induction on k, following the same pattern as Lemma 8.17 in Chapter 8.

Initial step ($k = 1$): Apply Lemma 9.14 with $T_0 = t_1$, initial radius $\rho_1 := \rho$, and the set Q_1. There exists $\mu_1 \geq \mu_1^\star$ such that the control $u^\star(\cdot,t) = u_{\mu_1}(\cdot)$ defined by (9.25) for $t \in [0, t_1 + \tau]$ yields $S(V(x,t)) = \mathbf{1}_{Q_1}(x)$ for all $t \in [t_1, t_1 + \tau]$. By the uniform bound (9.27), the state satisfies $\|V(\cdot,t)\|_{L^\infty} \leq \rho_2 := \rho + (b_F + \mu_1)/\lambda$ for all $t \in [0, t_1 + \tau]$.

Induction step: Suppose that after the $(k-1)$-th step, $\|V(\cdot, t_{k-1} + \tau)\|_{L^\infty} \leq \rho_k$. Apply Lemma 9.14 on the interval $[t_{k-1} + \tau,\ t_k + \tau]$ with initial radius ρ_k and $T_0 = t_k - t_{k-1} - \tau > 0$, for the set Q_k. There exists $\mu_k \geq \mu_k^\star(\rho_k, T_0)$ such that the control $u^\star(\cdot,t) = u_{\mu_k}(\cdot)$ for $t \in [t_{k-1} + \tau, t_k + \tau]$ yields $S(V(x,t)) = \mathbf{1}_{Q_k}(x)$ for all $t \in [t_k, t_k + \tau]$. The updated bound is

$$\rho_{k+1} = \rho_k + \frac{b_F + \mu_k}{\lambda}. \tag{9.34}$$

Note that each term $(b_F + \mu_k)/\lambda$ is independent of the duration $t_k - t_{k-1}$, thanks to the dissipative structure. Iterating up to $k = N$ and assembling the piecewise constant control over the intervals concludes the proof. $\square$

These two lemmas, combined with the dyadic partition of Lemma 9.5, provide all the ingredients for the proof of Theorem 9.7.

9.6 PROOF OF THE MAIN THEOREM

We now prove Theorem 9.7 by combining the technical lemmas developed in the previous section with the dyadic partition of Lemma 9.5.

Proof of Theorem 9.7. We combine the technical lemmas developed in the previous section with the dyadic partition of Lemma 9.5.

Step 1: Sequential control design. The dyadic partition $(Q_k, \tilde{e}_k)_{k=1}^{N_j}$ of Lemma 9.5 provides measurable cells $Q_k \subset \Omega$ with $|Q_k| > 0$. By hypothesis, the time instants $t_1 < t_2 < \cdots < t_{N_j}$ satisfy $t_1 \geq \tau$ and $t_{k+1} > t_k + \tau$. By Lemma 9.17, there exists a piecewise constant control $u^\star$ such that

$$S(V(x,t)) = \mathbf{1}_{Q_k}(x), \quad \text{a.e. } x \in \Omega, \quad \forall t \in [t_k, t_k + \tau], \tag{9.35}$$

for all $k = 1, \ldots, N_j$ and any initial condition with $\|V(\cdot, 0)\|_{L^\infty} \leq \rho$.

Step 2: Output extraction. During the time interval $[t_k, t_k + \tau]$, the output satisfies

$$y(s) = W\,[S(V(\cdot, s))] = W\,\mathbf{1}_{Q_k}, \quad \forall s \in [t_k, t_k + \tau]. \tag{9.36}$$

In particular, y is constant on each such interval, and the time average yields exactly

$$\hat{y}_k = |Q_k|^{-1/2} \frac{1}{\tau} \int_{t_k}^{t_k+\tau} y(s)\, ds = |Q_k|^{-1/2} W\,\mathbf{1}_{Q_k} = W\tilde{e}_k. \tag{9.37}$$

The image $W\tilde{e}_k$ is therefore recovered exactly: $W\tilde{e}_k = \hat{y}_k$. There is no approximation error.

Step 3: Error bound via Poincaré–Wirtinger. Since $W\tilde{e}_k$ is known exactly for all k, the estimator satisfies $\hat{W}_j = W \circ P_j$ where P_j is the orthogonal projection onto $E_j = \operatorname{span}(\tilde{e}_1, \ldots, \tilde{e}_{N_j})$. The error is therefore $\|W - \hat{W}_j\|_{\mathrm{op}} = \|W(I - P_j)\|_{\mathrm{op}}$. Factoring through the adjoint:

$$\|W(I - P_j)\|_{L^2 \to \mathcal{G}} = \|(I - P_j)W^*\|_{\mathcal{G} \to L^2} \leq \|I - P_j\|_{H^1 \to L^2}\, \|W^*\|_{\mathcal{G} \to H^1} \leq C_P\, h\, \|W^*\|_{\mathcal{G} \to H^1}, \tag{9.38}$$

where we used $\|T\|_{\mathrm{op}} = \|T^*\|_{\mathrm{op}}$ with $I - P_j$ self-adjoint, and $\|I - P_j\|_{H^1 \to L^2} \leq C_P\, h$ from the Poincaré–Wirtinger inequality (9.6) applied cell by cell. This establishes (9.11). $\square$

9.7 EXTENSION TO BOUNDED LINEAR OPERATORS

We now extend the results by adding a bounded spatial coupling. We set $\mathcal{A} = -\lambda I + \mathcal{A}_0$ with $\mathcal{A}_0 \in \mathcal{L}(L^\infty(\Omega))$. The system becomes

$$\begin{cases} \partial_t V = -\lambda V + \mathcal{A}_0 V + F(V, t) + u, \\ y = \mathcal{W}[S(V(\cdot, t))], \end{cases} \tag{9.39}$$

with the same output operator $\mathcal{W}$, saturation function S, and bounded perturbation F ($\|F(V,t)\|_{L^\infty} \le b_F$) as in system (9.3).

The key observation is that the pointwise argument of Lemma 9.14 extends when the coupling $\mathcal{A}_0$ is bounded. Indeed, the obstruction for *unbounded* coupling (e.g., the Laplacian) is that $|(\mathcal{A}_0 V)(x)|$ cannot be bounded uniformly over $\|V\|_{L^\infty} \le R$, so the coupling term cannot be dominated by any finite control amplitude (there will be a discussion on the unbounded case in section 9.9). When $\mathcal{A}_0$ is bounded, $|(\mathcal{A}_0 V)(x)| \le \|\mathcal{A}_0\|_{L^\infty \to L^\infty} \|V\|_{L^\infty}$ for a.e. $x \in \Omega$, and the coupling can be absorbed into the bounded perturbation.

We set $\Lambda_0 := \|\mathcal{A}_0\|_{L^\infty \to L^\infty}$. The semigroup $e^{t(-\lambda I + \mathcal{A}_0)} = e^{-\lambda t} e^{t\mathcal{A}_0}$ satisfies $\|e^{t\mathcal{A}}\|_{L^\infty \to L^\infty} \le e^{-(\lambda - \Lambda_0)t}$ (with $M_\infty = 1$, $\omega_\infty = \lambda - \Lambda_0$) whenever $\Lambda_0 < \lambda$. Define

$$\kappa := \frac{\Lambda_0}{\lambda - \Lambda_0}, \tag{9.40}$$

which is identical to the finite-dimensional formula of Chapter 8.

Assumption 9.18. *The operator $\mathcal{A}_0 \in \mathcal{L}(L^\infty(\Omega))$ satisfies $\kappa < 1$, equivalently $\Lambda_0 < \lambda/2$.*

Lemma 9.19 (Indicator function realization, bounded operator)**.** *Under Assumption 9.18, let $\rho > 0$ and $T_0 > 0$. Let $Q \subset \Omega$ be a measurable set with $|Q| > 0$. There exists $\mu^\star > 0$ such that for any $\mu \ge \mu^\star$, the constant-in-time control*

$$u_\mu(x) = \mu \mathbf{1}_Q(x) - \mu \mathbf{1}_{\Omega \setminus Q}(x) \tag{9.41}$$

yields, for any solution of (9.39) initialized with $\|V(\cdot,0)\|_{L^\infty(\Omega)} \le \rho$:

$$S(V(x,t)) = \mathbf{1}_Q(x), \quad \text{for a.e. } x \in \Omega, \quad \forall t \ge T_0. \tag{9.42}$$

The sufficient control amplitude is

$$\mu^\star = \frac{1}{1-\kappa} \left[\frac{\lambda(R+\rho)}{1 - e^{-\lambda T_0}} + (1+\kappa)\, b_F + \Lambda_0 \rho \right], \tag{9.43}$$

where $R = \max\{|\underline{R}|, \bar{R}\}$ and κ is defined by (9.40). For $\kappa = 0$ (i.e., $\mathcal{A}_0 = 0$), this reduces to the scalar formula of Lemma 9.14.

Moreover, $\|V(\cdot,t)\|_{L^\infty} \le \rho + (b_F + \mu)/(\lambda - \Lambda_0)$ for all $t \ge 0$.

Proof. Since $\Lambda_0 < \lambda$, Assumption 9.11 holds with $M_\infty = 1$ and $\omega_\infty = \lambda - \Lambda_0$. By Proposition 9.12, the solution satisfies the global bound

$$\|V(\cdot,t)\|_{L^\infty} \le R_\mu := \rho + \frac{b_F + \mu}{\lambda - \Lambda_0}, \quad \forall t \ge 0. \tag{9.44}$$

Upper saturation on Q. Since $\mathcal{A} = -\lambda I + \mathcal{A}_0$ is bounded, the mild solution is classical by Proposition 9.13, and the pointwise identity (9.24) holds. For a.e. $x \in Q$, the dynamics read

$$\partial_t V(x,t) = -\lambda V(x,t) + \underbrace{(\mathcal{A}_0 V)(x,t) + F(V,t)(x)} + \mu. \tag{9.45}$$

Since $|(\mathcal{A}_0 V)(x,t)| \le \Lambda_0 \|V(\cdot,t)\|_{L^\infty} \le \Lambda_0 R_\mu$ and $|F(V,t)(x)| \le b_F$:

$$\partial_t V(x,t) \ge -\lambda V(x,t) + \mu - \Lambda_0 R_\mu - b_F =: -\lambda V(x,t) + c_\mu. \tag{9.46}$$

Since $\Lambda_0 R_\mu = \Lambda_0 \rho + \frac{\Lambda_0}{\lambda - \Lambda_0}(b_F + \mu) = \Lambda_0 \rho + \kappa(b_F + \mu)$:

$$c_\mu = (1-\kappa)\mu - (1+\kappa)\, b_F - \Lambda_0 \rho. \tag{9.47}$$

Since $\kappa < 1$, we have $c_\mu \to +\infty$ as $\mu \to +\infty$. By the scalar comparison principle (as in the proof of Lemma 9.14) with $V(x,0) \ge -\rho$:

$$V(x,t) \ge -\rho\, e^{-\lambda t} + \frac{c_\mu}{\lambda}(1 - e^{-\lambda t}). \tag{9.48}$$

For $t \ge T_0$, this exceeds $\bar{R}$ provided

$$c_\mu \ge \frac{\lambda(\bar{R} + \rho)}{1 - e^{-\lambda T_0}}. \tag{9.49}$$

Moreover, since $c_\mu/\lambda > \bar{R}$, the equilibrium of the comparison ODE exceeds $\bar{R}$, so $V(x,t) \ge \bar{R}$ is maintained for all $t \ge T_0$ by the same dissipative argument as in Lemma 9.14. Hence $S(V(x,t)) = 1$ for a.e. $x \in Q$ and all $t \ge T_0$.

Lower saturation on $\Omega \setminus Q$. For a.e. $x \in \Omega \setminus Q$, the symmetric argument gives

$$\partial_t V(x,t) \le -\lambda V(x,t) - \mu + \Lambda_0 R_\mu + b_F = -\lambda V(x,t) - c_\mu, \tag{9.50}$$

and the comparison principle with $V(x,0) \le \rho$ yields $V(x,t) \le \underline{R}$ for all $t \ge T_0$ provided $c_\mu \ge \lambda(|\underline{R}| + \rho)/(1 - e^{-\lambda T_0})$, giving $S(V(x,t)) = 0$.

Taking the maximum of both thresholds and solving for μ yields (9.43). The uniform state bound follows from (9.44). □

Since the indicator function realization is *exact* (as in the scalar case), the sequential realization lemma, the main theorem, and its proof carry over verbatim from the scalar case. We state the result for completeness.

Theorem 9.20. *Under the assumptions of Table 9.1 and Assumption 9.18, the conclusions of Theorem 9.7 hold for system (9.39): the images $W\tilde{e}_k$ are recovered* exactly*, the estimator satisfies the same error bound (9.11), and the only source of error is the projection error controlled by the mesh width h.*

Proof. The proof is identical to that of Theorem 9.7. Lemma 9.19 provides exact indicator function realization $S(V(x,t)) = \mathbf{1}_Q(x)$ for a.e. $x \in \Omega$ and all $t \ge T_0$, with uniform-in-time L^∞ bounds that allow the inductive construction of Lemma 9.17 to proceed unchanged. The output extraction and error analysis (Steps 2–3 of the proof of Theorem 9.7) use only $S(V(\cdot,t)) = \mathbf{1}_{Q_k}$ exactly, which holds here as in the scalar case. □

Remark 9.21 (Comparison with the scalar and semigroup cases)**.** *The bounded operator extension occupies an intermediate position:*

- Scalar case *($\mathcal{A}_0 = 0$, $\kappa = 0$): No spatial coupling, no condition beyond exponential stability, minimal control amplitude $\mu^\star = b_F + \lambda(R+\rho)/(1 - e^{-\lambda T_0})$.*

- Bounded coupling *($\kappa < 1$): Spatial coupling via $\mathcal{A}_0 \in \mathcal{L}(L^\infty)$. Control amplitude scales as $\mu^\star \sim 1/(1-\kappa)$, diverging as $\kappa \to 1$. The condition $\kappa < 1$ is precisely what ensures that the resolvent $(-\mathcal{A})^{-1}$ preserves the sign of the control input, so the pointwise argument of Lemma 9.14 extends.*

- Unbounded operator *(Section 9.9): Spatial coupling is unbounded. The pointwise argument fails entirely, and one must resort to Green's function analysis with a structural resolution limit* $\operatorname{inrad}(Q) > \delta_0$.

9.8 KNOWN LINEAR PART $\mathcal{A}$

In this section, we extend the previous results to the case where the full linear operator $\mathcal{A}$ is known. Unlike the bounded-operator extension (Section 9.7), where the coupling $\mathcal{A}_0$ was unknown and the condition $\kappa < 1$ was required, the knowledge of $\mathcal{A}$ allows us to compute the steady-state map $\varphi \mapsto (-\mathcal{A})^{-1}B\varphi$ explicitly and to verify the controllability condition directly on $\operatorname{Im}((-\mathcal{A})^{-1}B)$. This is the infinite-dimensional analogue of the known-A extension of Section 8.6.

We consider the system

$$\begin{cases} \partial_t V = \mathcal{A}V + F(V,t) + Bu, \\ y = \mathcal{W}[S(V(\cdot,t))], \end{cases} \tag{9.51}$$

where F and $\mathcal{W}$ are as in Table 9.1, $B \in \mathcal{L}(U, L^\infty(\Omega))$ is a known bounded control operator with U a Banach space, and $\mathcal{A}$ satisfies the following:

Assumption 9.22. *The operator $\mathcal{A} : D(\mathcal{A}) \subset L^\infty(\Omega) \to L^\infty(\Omega)$ is known and generates an exponentially stable C_0-semigroup $(e^{t\mathcal{A}})_{t\geq 0}$ on $L^\infty(\Omega)$, with constants $M_\infty \geq 1$ and $\omega_\infty > 0$ such that $\|e^{t\mathcal{A}}\|_{L^\infty \to L^\infty} \leq M_\infty e^{-\omega_\infty t}$ for all $t \geq 0$. In particular, $0 \in \rho(\mathcal{A})$ and $(-\mathcal{A})^{-1} = \int_0^\infty e^{t\mathcal{A}}\,dt \in \mathcal{L}(L^\infty(\Omega))$.*

Since both $\mathcal{A}$ and B are known, the resolvent $(-\mathcal{A})^{-1}B$ is computable. Given a control direction $v \in U$, the Duhamel formula with constant control $u(t) = \mu v$ for system (9.51) yields

$$V(t) = \underbrace{\mu\tilde{\varphi}}_{\text{steady state}} + \underbrace{e^{t\mathcal{A}}(V(0) - \mu\tilde{\varphi})}_{r_1(t)} + \underbrace{\int_0^t e^{(t-s)\mathcal{A}} F(V(s),s)\,ds}_{r_2(t)}, \tag{9.52}$$

where $\tilde{\varphi} := (-\mathcal{A})^{-1}Bv \in L^\infty(\Omega)$ is the steady-state target (the particular solution of $\mathcal{A}V + Bv = 0$). The transient satisfies $\|r_1(t)\|_{L^\infty} \leq M_\infty e^{-\omega_\infty t}(\rho + \mu\|\tilde{\varphi}\|_{L^\infty})$, and the nonlinear residual $\|r_2(t)\|_{L^\infty} \leq M_\infty b_F/\omega_\infty$ uniformly in t.

Since both $\mathcal{A}$ and B are known, the sign structure of $\tilde{\varphi} = (-\mathcal{A})^{-1}Bv$ is directly checkable for each control direction $v \in U$. The controllability condition is therefore imposed on $\operatorname{Im}((-\mathcal{A})^{-1}B)$: this replaces the condition $\kappa < 1$, which was a sufficient condition ensuring that $(-\mathcal{A})^{-1}Bv$ retains the sign of Bv when $\mathcal{A}$ is unknown. We impose the following condition, which is the infinite-dimensional analogue of Assumption 8.20:

Assumption 9.23. *There exists $v \in U$ with $\|v\|_U = 1$ and a measurable set $Q \subset \Omega$ with $|Q| > 0$ such that, setting $\tilde{\varphi} := (-\mathcal{A})^{-1}Bv$,*

$$\underline{\tilde{\varphi}} := \min\Big(\operatorname{ess\,inf}_{x \in Q} \tilde{\varphi}(x),\ \operatorname{ess\,inf}_{x \in \Omega \setminus Q} \big(-\tilde{\varphi}(x)\big) \Big) > 0. \tag{9.53}$$

In other words, $\tilde{\varphi}(x) \geq \underline{\tilde{\varphi}} > 0$ for a.e. $x \in Q$ and $\tilde{\varphi}(x) \leq -\underline{\tilde{\varphi}} < 0$ for a.e. $x \in \Omega \setminus Q$.

Remark 9.24. *In the bounded-operator case (Section 9.7), the operator was decomposed as $\mathcal{A} = -\lambda I + \mathcal{A}_0$ with $B = \mathrm{Id}$ and $Bv = \mathbf{1}_Q - \mathbf{1}_{\Omega \setminus Q}$, giving $\tilde{\varphi} = (-\mathcal{A})^{-1}Bv$. The condition $\kappa < 1$ was a sufficient condition ensuring that $\tilde{\varphi}$ retains the sign of Bv (positive on Q, negative on $\Omega \setminus Q$) when $\mathcal{A}_0$ is unknown. When $\mathcal{A}$ is known, one computes $\tilde{\varphi} = (-\mathcal{A})^{-1}Bv$ directly and checks Assumption 9.23 without any smallness condition.*

We consider two subcases: a general case where only Assumptions 9.22 and 9.23 hold, and one where we additionally have controllability of the pair $(\mathcal{A}, B)$.

9.8.1 *Known linear part: General Case*

The following is the counterpart of Lemma 8.23 in infinite dimension:

Lemma 9.25 (Indicator realization, known $\mathcal{A}$)**.** *Under Assumptions 9.22 and 9.23, let $\rho > 0$. Define the minimal steering time*

$$T_{\min} := \frac{1}{\omega_\infty} \ln\left(\frac{M_\infty \|\tilde{\varphi}\|_{L^\infty}}{\underline{\tilde{\varphi}}} \right). \tag{9.54}$$

For any $T_0 > T_{\min}$, let $\gamma(T_0) := \underline{\tilde{\varphi}} - M_\infty e^{-\omega_\infty T_0} \|\tilde{\varphi}\|_{L^\infty} > 0$. There exists $\mu^\star > 0$ such that for all $\mu \geq \mu^\star$, the constant-in-time control $u = \mu v$ yields, for any mild solution of (9.51) with $\|V(\cdot, 0)\|_{L^\infty} \leq \rho$:

$$S(V(x,t)) = \mathbf{1}_Q(x), \quad \text{for a.e. } x \in \Omega, \quad \forall t \geq T_0. \tag{9.55}$$

The sufficient control amplitude is

$$\mu^\star = \frac{R + M_\infty \rho + M_\infty b_F / \omega_\infty}{\gamma(T_0)}, \tag{9.56}$$

where $R = \max\{|\underline{R}|, \bar{R}\}$. Moreover, $\|V(\cdot,t)\|_{L^\infty} \leq \mu \|\tilde{\varphi}\|_{L^\infty}(1 + M_\infty) + M_\infty \rho + M_\infty b_F/\omega_\infty$ for all $t \geq 0$.

Proof. From decomposition (9.52), rewriting $V(t) = \mu(I - e^{t\mathcal{A}})\tilde{\varphi} + e^{t\mathcal{A}}V(0) + r_2(t)$, we estimate pointwise. For a.e. $x \in Q$ and $t \geq T_0$:

$$V(x,t) \geq \mu\big(\tilde{\varphi}(x) - M_\infty e^{-\omega_\infty T_0} \|\tilde{\varphi}\|_{L^\infty}\big) - M_\infty \rho - \frac{M_\infty b_F}{\omega_\infty} \geq \mu\, \gamma(T_0) - M_\infty \rho - \frac{M_\infty b_F}{\omega_\infty}, \tag{9.57}$$

using $\tilde{\varphi}(x) \geq \underline{\tilde{\varphi}}$ and $\|e^{t\mathcal{A}}\|_{L^\infty \to L^\infty} \leq M_\infty e^{-\omega_\infty t}$. The coefficient $\gamma(T_0) > 0$ precisely when $T_0 > T_{\min}$, and the right-hand side exceeds R for $\mu \geq \mu^\star$. The symmetric argument with $\tilde{\varphi}(x) \leq -\underline{\tilde{\varphi}}$ on $\Omega \setminus Q$ gives $V(x,t) \leq -R$ there. As $e^{-\omega_\infty t}$ is decreasing, both inequalities persist for all $t \geq T_0$, so $S(V(\cdot,t)) = \mathbf{1}_Q$ permanently. The state bound follows from the triangle inequality applied to (9.52). □

9.8.2 *Known linear part: Controllable Case*

The minimal time $T_{\min}$ in Lemma 9.25 arises because a constant control produces a transient of order μ that must decay before the sign structure locks. As in the finite-dimensional case (Lemma 8.24), this constraint can be eliminated under a controllability assumption on the pair $(\mathcal{A}, B)$.

Assumption 9.26. *The pair $(\mathcal{A}, B)$ is exactly controllable in time $T_0 > 0$: for every $\psi \in L^\infty(\Omega)$, there exists a control $u_0 \in L^\infty([0, T_0]; U)$ such that*

$$\int_0^{T_0} e^{(T_0-s)\mathcal{A}} B u_0(s)\, ds = \psi. \tag{9.58}$$

Remark 9.27. *When U and $L^2(\Omega)$ are Hilbert spaces and $B \in \mathcal{L}(U, L^2(\Omega))$, exact controllability in time T_0 is equivalent to the invertibility of the controllability Gramian*

$$\mathcal{G}_B(T_0) := \int_0^{T_0} e^{s\mathcal{A}} B B^* e^{s\mathcal{A}^*}\, ds, \tag{9.59}$$

and the minimum-norm steering control to reach ψ is given by $u_0(s) = B^* e^{(T_0-s)\mathcal{A}^*} \mathcal{G}_B(T_0)^{-1}\psi$, which is the infinite-dimensional analogue of the Gramian-based control in Lemma 8.24. When $B = \mathrm{Id}$ and $\mathcal{A}$ generates a C_0-group, Assumption 9.26 is automatically satisfied. Indeed, the explicit control $u_0(s) = \frac{1}{T_0} e^{-(T_0-s)\mathcal{A}}\psi$ achieves $\int_0^{T_0} e^{(T_0-s)\mathcal{A}} u_0(s)\, ds = \psi$ for any $T_0 > 0$.

Lemma 9.28 (Indicator realization, known $\mathcal{A}$, any T_0)**.** *Under Assumptions 9.22, 9.23, and 9.26, for any $T_0 > 0$ there exists $\mu^\star > 0$ and a bounded time-varying control $u_\mu \in L^\infty([0, +\infty); U)$ such that, for any mild solution of (9.51) with $\|V(\cdot, 0)\|_{L^\infty} \le \rho$:*

$$S(V(x,t)) = \mathbf{1}_Q(x), \quad \text{for a.e. } x \in \Omega, \quad \forall t \ge T_0. \tag{9.60}$$

The sufficient control amplitude is $\mu^\star = (R + M_\infty \rho + M_\infty b_F/\omega_\infty)/\underline{\tilde{\varphi}}$.

Proof. By Assumption 9.26, there exists $u_0 \in L^\infty([0, T_0]; U)$ such that $\int_0^{T_0} e^{(T_0-s)\mathcal{A}} B u_0(s)\, ds = \tilde{\varphi}$. Define the control

$$u(t) = \begin{cases} \mu\, u_0(t), & t \in [0, T_0], \\ \mu\, v, & t > T_0, \end{cases} \tag{9.61}$$

where $v \in U$ is the control direction from Assumption 9.23. By the Duhamel formula at $t = T_0$:

$$V(T_0) = \mu \tilde{\varphi} + r, \qquad r := e^{T_0 \mathcal{A}} V(0) + \int_0^{T_0} e^{(T_0-s)\mathcal{A}} F(V(s), s)\, ds, \tag{9.62}$$

where $\|r\|_{L^\infty} \le M_\infty \rho + M_\infty b_F/\omega_\infty$ is independent of μ. For $t \ge T_0$, the constant control $Bu = \mu B v$ and the identity $\tilde{\varphi} = (-\mathcal{A})^{-1} B v$ give $V(T_0) + \mu \mathcal{A}^{-1} B v = r$, so the Duhamel formula yields

$$V(t) = \mu \tilde{\varphi} + e^{(t-T_0)\mathcal{A}} r + \int_{T_0}^{t} e^{(t-s)\mathcal{A}} F(V(s), s)\, ds. \tag{9.63}$$

Both residual terms are bounded independently of μ and decay exponentially. For a.e. $x \in Q$: $V(x,t) \ge \mu \underline{\tilde{\varphi}} - M_\infty \rho - M_\infty b_F/\omega_\infty \ge R$ for $\mu \ge \mu^\star$. The symmetric argument on $\Omega \setminus Q$ gives $V(x,t) \le -R$, so $S(V(\cdot, t)) = \mathbf{1}_Q$ for all $t \ge T_0$. $\square$

9.8.3 *Conclusion*

The two lemmas above are the infinite-dimensional analogues of Lemmas 8.23 and 8.24. The sign-structure condition is imposed on $\mathrm{Im}((-\mathcal{A})^{-1}B)$ instead of on $\mathrm{Im}(B)$ alone, and since both $\mathcal{A}$ and B are known, this condition is directly checkable. Under these conditions, the sequential saturation protocol and the parameter recovery of Theorem 9.20 carry over to system (9.51), with the condition $\kappa < 1$ of Assumption 9.18 replaced by the sign condition of Assumption 9.23.

9.9 PERSPECTIVES: UNBOUNDED SPATIAL COUPLING

All the results established in this chapter share the same underlying mechanism, which can be traced back to the Duhamel formula. Under a constant control $u \in L^\infty(\Omega)$, the mild solution of system (9.39) satisfies

$$V(t) = e^{t\mathcal{A}}V(0) + \int_0^t e^{(t-s)\mathcal{A}}F(V(s),s)\,ds + \int_0^t e^{(t-s)\mathcal{A}}u\,ds. \tag{9.64}$$

The last term converges, as $t \to \infty$, to $(-\mathcal{A})^{-1}u$ which is the steady state of $\mathcal{A}V + u = 0$. The first two terms are bounded independently of μ. Consequently, for large enough μ, the behavior of $V(x,t)$ at large times is dominated by $V_\mu := (-\mathcal{A})^{-1}(\mu u)$, and the condition $S(V(x,t)) = \mathbf{1}_Q(x)$ reduces to whether $(-\mathcal{A})^{-1}u$ is positive on Q and negative on $\Omega \setminus Q$.

In the bounded-coupling case, $\mathcal{A} = -\lambda I + \mathcal{A}_0$. Under Assumption 9.18, $\Lambda_0 < \lambda$, so the resolvent admits the absolutely convergent Neumann series

$$(-\mathcal{A})^{-1} = \frac{1}{\lambda}\sum_{k=0}^{\infty}\left(\frac{\mathcal{A}_0}{\lambda}\right)^k, \tag{9.65}$$

whose dominant term $\frac{1}{\lambda}I$ maps $u = \mathbf{1}_Q - \mathbf{1}_{\Omega\setminus Q}$ to $\frac{1}{\lambda}u$, which already has the correct sign. The remainder term $\frac{1}{\lambda}\sum_{k\geq 1}(\mathcal{A}_0/\lambda)^k$ has L^∞-operator norm at most $\frac{\Lambda_0}{\lambda(\lambda-\Lambda_0)} = \frac{\kappa}{\lambda}$, so the remainder is bounded by κ relative to the leading term $1/\lambda$. Hence for $\kappa < 1$ the sign structure of u is preserved by $(-\mathcal{A})^{-1}$, pointwise and uniformly over Ω. This is the core reason that exact indicator realization is available in all three bounded cases (scalar, bounded coupling, known $\mathcal{A}$).

When $\mathcal{A}$ is an unbounded operator, we can no longer exploit the above structure, however we can still utilize Duhamel's formula. When the unbounded operator $\mathcal{A}$, which can be assumed to be a diffusion operator, generates a strongly continuous exponentially stable semigroup on $L^\infty(\Omega)$, the resolvent $(-\mathcal{A})^{-1}$ is still a well-defined bounded operator on $L^\infty(\Omega)$. Supposing it admits a Green's function $G(x,y)$, one would have then:

$$V_\mu(x) = \left[(-\mathcal{A})^{-1}u\right](x) = \int_\Omega G(x,y)\,u(y)\,dy. \tag{9.66}$$

With the piecewise-constant control $u = \mu_+ \mathbf{1}_Q - \mu_- \mathbf{1}_{\Omega\setminus Q}$, this decomposes into near-field and far-field contributions:

$$V_\mu(x) = \mu_+ \underbrace{\int_Q G(x,y)\,dy}_{a(x)} - \mu_- \underbrace{\int_{\Omega\setminus Q} G(x,y)\,dy}_{b(x)}. \tag{9.67}$$

Assuming $G > 0$ a.e. (which is a reasonable assumption for many diffusion operators [74]), both $a(x)$ and $b(x)$ are strictly positive at every $x \in \Omega$. The sign of $V_\mu(x)$ is therefore not determined by $u(x)$ alone, but by the competition between the near-field integral $a(x)$ and the far-field integral $b(x)$.

Remark 9.29. *A concrete example of such operators would be $\mathcal{A} = \nabla \cdot (a(x)\nabla) - \lambda_0$ with uniformly elliptic coefficients ($a_-|\xi|^2 \le \xi^\top a(x)\xi \le a_+|\xi|^2$) and $\lambda_0 > 0$, subject to Dirichlet or Neumann boundary conditions, then $\mathcal{A}$ generates an analytic, exponentially stable, positive semigroup on $L^\infty(\Omega)$[134, Chapter 7], [74, Chapter 6] and the resolvent $(-\mathcal{A})^{-1}$ is a well-defined bounded operator on $L^\infty(\Omega)$ by the Hille-Yosida theorem. The Green's function of $(-\mathcal{A})^{-1}$ satisfies two-sided Gaussian bounds $G(x,y) \asymp e^{-\gamma|x-y|}$ with $\gamma = \sqrt{\lambda_0/a_+}$; see [95, Theorem 1.4] for the upper bound and [57, Chapter 4] for two-sided estimates via heat kernel integration, and [91, Chapter 4] for the classical potential-theoretic approach.*

For a point x lying deep inside Q, at distance $\mathrm{dist}(x, \partial Q) > \delta$ from the boundary, the near-field integral $a(x)$ concentrates contributions from the ball $B_\delta(x) \subset Q$, while $b(x)$ collects only far-away points. If G decays sufficiently fast with distance, for instance, with Gaussian-type bounds $G(x,y) \sim e^{-\gamma|x-y|}$, then $a(x)$ grows with δ and $b(x)$ shrinks. There exists a critical width $\delta_0 > 0$, set by the spatial decay rate γ of G (roughly $\delta_0 \sim 1/\gamma$), below which the near-field and far-field contributions are comparable and the sign of $V_\mu(x)$ cannot be guaranteed. For $\mathrm{dist}(x, \partial Q) > \delta_0$, one can choose μ_+, μ_- large enough to ensure $V_\mu(x) \ge \bar{R}$ inside Q and $V_\mu(x) \le \underline{R}$ outside. This is illustrated in Figure 9.1. Returning then to the Duhamel formula through appropriate choice of μ_+, μ_-, one can force $V_\mu(x)$ to have the correct sign structure for all x with $\mathrm{dist}(x, \partial Q) > \delta_0$, ensuring that $S(V(\cdot,t)) = \mathbf{1}_{\mathrm{dist}(x,\partial Q) > \delta_0 \cap x \in Q} + r(\delta_0)$ after a certain time $T_0 > 0$ with a certain error $|r(\delta_0)| \le C_G \delta_0$ which will depend on the nature of the Green's function G.

The condition on the Green's function (two-sided Gaussian bounds) is precisely the assumption needed to make this heuristic rigorous; it quantifies $a(x)$ from below and $b(x)$ from above, yielding an explicit formula for δ_0 and explicit thresholds $\mu_+^\star, \mu_-^\star$ on the control amplitude. Once the approximation of $S(V(\cdot,t))$ by $\mathbf{1}_{\mathrm{dist}(x,\partial Q) > \delta_0 \cap x \in Q}$ is established, one could then obtain the same type of results as in Theorem 9.20 with an error term depending on the chosen $\epsilon > 0$ and a term depending on δ_0 that depends on the nature of the Green's function G and cannot be compensated.

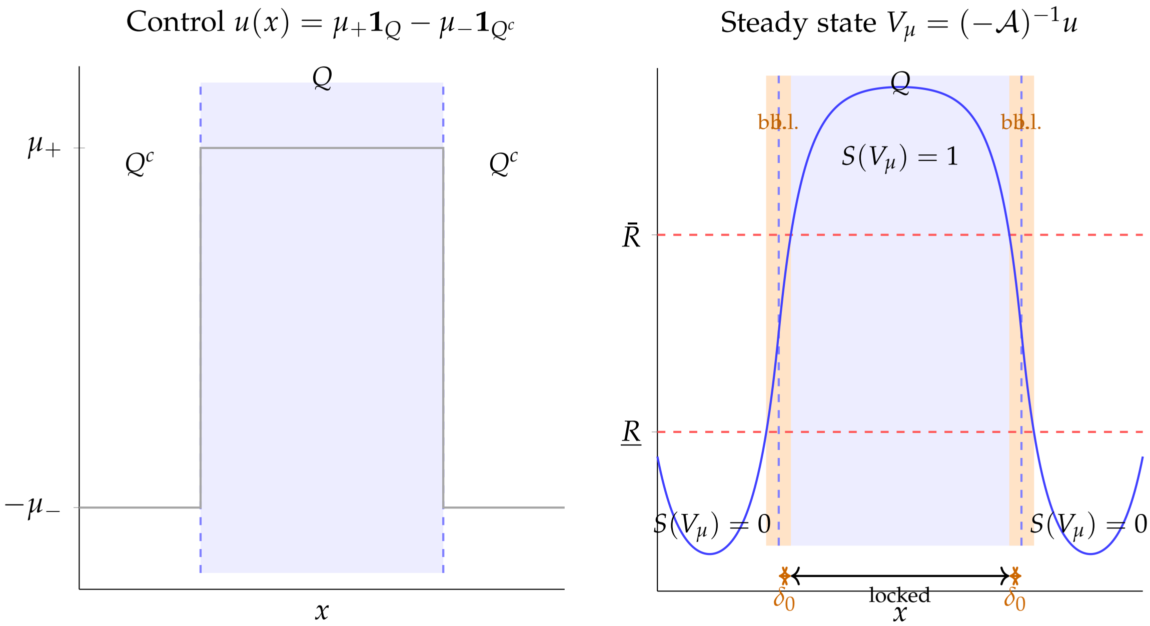


Figure 9.1: Left: the piecewise-constant control $u(x) = \mu_+ \mathbf{1}_Q(x) - \mu_- \mathbf{1}_{Q^c}(x)$. Right: the resulting steady state $V_\mu(x) = (-\mathcal{A})^{-1}u$. The Green's function mixes near-field and far-field contributions near ∂Q, producing boundary layers (orange, b.l.) of width δ_0. In the locked region $(\text{dist}(x, \partial Q) > \delta_0)$, a sufficiently large amplitude μ enforces $S(V_\mu) = 1$ inside Q and $S(V_\mu) = 0$ outside.

10

TWO-LAYER NETWORKS OF FINITE DIMENSIONAL SYSTEMS

We consider a two-population network of neural mass models with saturated interconnections inspired by Wilson–Cowan equations, in which only the first population is observed and all connection weights are unknown. Offline parameter estimation methods struggle for this class of models as they typically rely on passive observation. Furthermore, if the inter-population coupling is linear, certain connection parameters are structurally unidentifiable. Saturation resolves this ambiguity by building on the results of Chapter 8. A five-stage protocol is proposed that uses controlled inputs to steer the saturation outputs through prescribed values and isolate each block of unknown parameters in turn. We show that all unknown parameters are identified exactly in finite time. The procedure is inherently online: each stage constructs its inputs from the estimates produced by the previous one. Numerical simulations are presented to illustrate the method.

10.1 INTRODUCTION

The application of control theory to neuroscience is motivated by a broad range of clinical and scientific objectives, from brain health monitoring [149] and brain-computer interfaces [150] to closed-loop neuromodulation for Parkinson's disease [62, 47]. All of these require accurate knowledge of the model parameters, which must therefore be estimated from measurements.

We focus on neural mass models based on the Wilson–Cowan framework [168], which track population-averaged activity and are a standard tool for studying macroscopic cortical dynamics [61, 36]. Parameter uncertainty in neural mass models is one of their most inherent drawbacks [51, 20]. Furthermore, parameters such as synaptic gains may drift over time [150], and offline methods cannot adapt the experimental inputs as estimation proceeds [128].

Since the state is only partially measured and the parameters are unknown, adaptive observers are a natural tool [69]. They have been applied to neural mass models in several settings: [44] for a conductance-based model, [40, 41] for a delayed neural fields model [167, 3], and [137] for a Jansen–Rit model [106], typically under dissipativity, persistent excitation, or strong structural assumptions on the system.

This chapter takes a different approach. Rather than relying on properties of the uncontrolled trajectory, we design the experimental inputs explicitly to force identifiability at each stage. By steering the saturation to prescribed values, we can lift the identification ambiguity and effectively retrieve the parameters by using a mix of an exact differentiator and least-squares estimators. The estimation is necessarily online and recursive, as each stage uses the estimates of the previous one to construct its inputs. The main result is that all unknown parameters of a general two-layer system are recovered exactly in finite time. The precise statement and the five-stage protocol are given in Section 10.2.

10.2 MAIN RESULT

NOTATIONS We follow the notation of Chapter 8. 1. For any two integers $p < q$, $[\![p, q]\!] = \{i \in \mathbb{Z},\ p \leq i \leq q\}$.
2. The vector $\mathbf{1}_d \in \mathbb{R}^d$ has all entries equal to 1.
3. Given a function $g : \mathbb{R} \to \mathbb{R}$ and a vector $z \in \mathbb{R}^n$, $g(z) \in \mathbb{R}^n$ denotes the vector $(g(z_1), \ldots, g(z_n))$.
4. For $L > 0$ and $\Omega \subset \mathbb{R}^n$, $C^{1,L}(\mathbb{R}, \Omega)$ denotes the set of C^1 trajectories taking values in Ω whose derivative is L-Lipschitz.
5. $\|\cdot\|_\infty$ denotes the max norm.
6. The function $S : \mathbb{R} \to \mathbb{R}$ is a saturation satisfying (H4) below. We set $R := \max\{|\underline{R}|, |\bar{R}|\}$.
7. For $\epsilon > 0$, $\Omega_{S,\epsilon} := [\underline{R} + \epsilon,\ \bar{R} - \epsilon]$ is the strict interior of the invertible region of S.
8. For matrices $M_1 \in \mathbb{R}^{n \times p}$, $M_2 \in \mathbb{R}^{n \times q}$ and a vector $v \in \mathbb{R}^n$, $[M_1 \mid M_2 \mid v]$ denotes the horizontal concatenation $[M_1,\ M_2,\ v] \in \mathbb{R}^{n \times (p+q+1)}$.

Consider the following two-population system,

$$\begin{cases} \dot{V}_1 = A_1 V_1 + W_{11} S(V_1) + W_{12} S(V_2) + u_1 + h_1, \\ \dot{V}_2 = A_2 V_2 + W_{21} S(V_1) + W_{22} S(V_2) + u_2 + h_2, \\ y = V_1, \end{cases} \tag{10.1}$$

where $V_i \in \mathbb{R}^{n_i}$, $u_i \in \mathbb{R}^{n_i}$, $y \in \mathbb{R}^{n_1}$, and only the first population is measured. The matrices A_i, W_{ij} and the vectors h_i are all unknown except where specified below.

ASSUMPTIONS.

(H1) A_1 is unknown and Hurwitz, with $\|A_1\| \leq \bar{A}_1$ for some known constant $\bar{A}_1 > 0$; h_1 has a known bound $\bar{h}_1$.

(H2) A_2 is known and Hurwitz, and h_2 has a known bound $\bar{h}_2$.

(H3) All four blocks W_{ij} are unknown with a common known bound $\bar{W}$. We assume $n_1 \geq n_2$; the block $W_{12} \in \mathbb{R}^{n_1 \times n_2}$ has rank n_2, so its left inverse $W_{12}^\dagger = (W_{12}^\top W_{12})^{-1} W_{12}^\top$ satisfies $W_{12}^\dagger W_{12} = I_{n_2}$.

(H4) $S : \mathbb{R} \to \mathbb{R}$ is a known, globally Lipschitz-continuous, piecewise C^1 function with known saturation levels $\underline{R} < \bar{R}$ satisfying

$$x \leq \underline{R} \ \Rightarrow\ S(x) = 0, \qquad x \geq \bar{R} \ \Rightarrow\ S(x) = 1,$$

and strictly increasing on $[\underline{R}, \bar{R}]$; in particular $\|S\|_\infty = 1$ and S^{-1} is well defined and Lipschitz on $(0, 1)$.

The main result of this chapter is to show that all unknown parameters of (10.1) can be estimated in finite time. Before introducing the main theorem we present some heuristics

10.2.1 *Heuristics*

We first outline the estimation protocol, which proceeds in five steps:

1. *Disconnect V_2:* a large constant input $u_2 = -\mu \mathbf{1}_{n_2}$ drives $S(V_2)$ to zero, reducing the first equation to $\dot{V}_1 = A_1 V_1 + W_{11} S(V_1) + u_1 + h_1$.
2. *Identify $[A_1 \mid h_1]$:* steer V_1 to track a PE reference $\gamma^* < \underline{R}$, so $S(V_1) = 0$ and the equation becomes $\dot{V}_1 - u_1 = [A_1 \mid h_1]\Phi_1$; a super-twisting observer supplies $\dot{V}_1 - u_1$ exactly, and least squares recovers $[A_1 \mid h_1]$.
3. *Identify W_{11}:* cycle $S(V_1)$ through $e_1, \ldots, e_{n_1}$ following the approach of Chapter 8; on each window the residual $\dot{V}_1 - u_1 - A_1 V_1 - h_1 = W_{11} e_i$ gives one column of W_{11} by averaging.
4. *Identify W_{12}:* likewise cycle $S(V_2)$ through $e_1, \ldots, e_{n_2}$; the residual $\phi = W_{12} e_j$ recovers W_{12} column by column.
5. *Recover V_2 and identify the second layer:* invert S to form $\hat{V}_2 = S^{-1}(W_{12}^{\dagger}\phi)$, differentiate $\hat{V}_2$ with a second super-twisting observer, and recover $[W_{21} \mid W_{22} \mid h_2]$ by least squares.

The protocol unfolds over a time sequence $0 < T_0 < T_0' < T_1 < T_2 < T_3 < T_4 < T_5$, which is freely chosen with sufficiently large interval lengths; Theorem 10.1 makes this precise. These labels are used throughout the description below. The observers and estimators used in each step are as follows.

A super-twisting observer runs continuously to supply $\dot{V}_1 - u_1$. For each $j \in [\![1, n_1]\!]$,

$$\begin{cases} \dot{\hat{z}}_{1,1,j} = u_{1,j} + \hat{z}_{1,2,j} + \lambda_1 \lceil y_j - \hat{z}_{1,1,j} \rfloor^{1/2}, \\ \dot{\hat{z}}_{1,2,j} \in \lambda_2 \operatorname{sign}(y_j - \hat{z}_{1,1,j}), \end{cases}$$

with $\hat{z}_{1,2} := (\hat{z}_{1,2,1}, \ldots, \hat{z}_{1,2,n_1})$. For gains (λ_1, λ_2) large enough relative to the bound on $\ddot{V}_1 - \dot{u}_1$, one has $\hat{z}_{1,2}(t) = \dot{V}_1(t) - u_1(t)$ for all $t \geq T_0$.

Once $\hat{z}_{1,2}$ is available (i.e. for $t \geq T_0$) and $S(V_1) = S(V_2) = 0$, the first regressor and its estimator are

$$\Phi_1(t) := \begin{bmatrix} y(t) \\ 1 \end{bmatrix}, \qquad [\hat{A}_1 \mid \hat{h}_1] := \left(\int_{T_0}^{T_0'} \hat{z}_{1,2}\, \Phi_1^\top \right) \left(\int_{T_0}^{T_0'} \Phi_1 \Phi_1^\top \right)^{-1},$$

where $[T_0, T_0']$ is the feedback-tracking window on which $V_1 \approx \gamma^*$ and $S(V_1) = 0$. Given $\hat{A}_1$ and $\hat{h}_1$, the residual $\phi_1(t) := \hat{z}_{1,2}(t) - \hat{A}_1 y(t) - \hat{h}_1$ equals $W_{11} S(V_1(t))$, and the estimator

$$\hat{W}_{11} := \frac{1}{\tau} \sum_{i=1}^{n_1} \int_{s_i}^{s_i+\tau} \phi_1\, e_i^\top,$$

where $[s_i, s_i+\tau] \subset [T_0', T_1]$ are disjoint windows on which $S(V_1) = e_i$, recovers W_{11}. Given $\hat{A}_1$, $\hat{h}_1$, $\hat{W}_{11}$, the residual

$$\phi(t) := \hat{z}_{1,2}(t) - \hat{A}_1 y(t) - \hat{W}_{11} S(y(t)) - \hat{h}_1$$

equals $W_{12} S(V_2(t))$ identically. When $S(V_2) = e_j$ on disjoint windows $[t_j, t_j+\tau] \subset [T_1, T_2]$,

$$\hat{W}_{12} := \frac{1}{\tau} \sum_{j=1}^{n_2} \int_{t_j}^{t_j+\tau} \phi\, e_j^\top$$

recovers W_{12}. With $\hat{W}_{12}$ available, the estimate of V_2 is defined for $t \geq T_3$ by

$$\hat{V}_2(t) := S^{-1}\big(\hat{W}_{12}^\dagger\, \phi(t)\big),$$

which is valid whenever $\hat{W}_{12}^\dagger\, \phi(t) \in (0,1)^{n_2}$, i.e. when $V_2(t) \in \Omega_{S,\delta}^{n_2}$ for some $\delta > 0$. A second super-twisting observer, now tracking $\hat{V}_2$, is activated from T_3 onward. For each $j \in [\![1, n_2]\!]$,

$$\begin{cases} \dot{\hat{z}}_{2,1,j} = \mathbf{1}_{t \geq T_3}(t)\big[u_{2,j} + \hat{z}_{2,2,j} + \lambda_1 \lceil \hat{V}_{2,j} - \hat{z}_{2,1,j} \rfloor^{1/2}\big], \\ \dot{\hat{z}}_{2,2,j} \in \mathbf{1}_{t \geq T_3}(t)\, \lambda_2 \operatorname{sign}(\hat{V}_{2,j} - \hat{z}_{2,1,j}), \end{cases}$$

with $\hat{z}_{2,2} := (\hat{z}_{2,2,1}, \ldots, \hat{z}_{2,2,n_2})$. On the sliding surface, $\hat{z}_{2,2,j}(t) = \dot{\hat{V}}_{2,j}(t) - u_{2,j}(t)$. Once $\hat{V}_2$ and $\hat{z}_{2,2}$ are both available on $[T_4, T_5] \subset [T_3, T_5]$, the second-layer regressor and estimator are

$$\hat{\Phi}_2(t) := \begin{bmatrix} S(y(t)) \\ S(\hat{V}_2(t)) \\ 1 \end{bmatrix}, \qquad [\hat{W}_{21} \mid \hat{W}_{22} \mid \hat{h}_2] := \Big(\int_{T_4}^{T_5} (\hat{z}_{2,2} - A_2 \hat{V}_2) \hat{\Phi}_2^\top\Big)\Big(\int_{T_4}^{T_5} \hat{\Phi}_2 \hat{\Phi}_2^\top\Big)^{-1}.$$

10.2.2 *Main Theorem*

Consider a time sequence $0 < T_0 < T_0' < T_1 < T_2 < T_3 < T_4 < T_5$, a $\tau > 0$ and gains $\lambda_1, \lambda_2 > 0$. The complete estimation protocol, gathering the above in the order of activation, is:

$$\begin{cases} \dot{\hat{z}}_{1,1,j} = u_{1,j} + \hat{z}_{1,2,j} + \lambda_1 \lceil y_j - \hat{z}_{1,1,j} \rfloor^{1/2}, \\ \dot{\hat{z}}_{1,2,j} \in \lambda_2 \operatorname{sign}(y_j - \hat{z}_{1,1,j}), \end{cases} \qquad j \in [\![1, n_1]\!], \tag{10.2a}$$

$$\Phi_1 := \begin{bmatrix} y \\ 1 \end{bmatrix}, \qquad [\hat{A}_1 \mid \hat{h}_1] := \Big(\int_{T_0}^{T_0'} \hat{z}_{1,2}\, \Phi_1^\top\Big)\Big(\int_{T_0}^{T_0'} \Phi_1 \Phi_1^\top\Big)^{-1}, \tag{10.2b}$$

$$\hat{W}_{11} := \frac{1}{\tau} \sum_{i=1}^{n_1} \int_{s_i}^{s_i+\tau} (\hat{z}_{1,2} - \hat{A}_1 y - \hat{h}_1)\, e_i^\top, \tag{10.2c}$$

$$\phi := \hat{z}_{1,2} - \hat{A}_1 y - \hat{W}_{11} S(y) - \hat{h}_1, \tag{10.2d}$$

$$\hat{W}_{12} := \frac{1}{\tau} \sum_{j=1}^{n_2} \int_{t_j}^{t_j+\tau} \phi\, e_j^\top, \tag{10.2e}$$

$$\hat{V}_2 := S^{-1}\big(\hat{W}_{12}^\dagger\, \phi\big), \tag{10.2f}$$

$$\begin{cases} \dot{\hat{z}}_{2,1,j} = \mathbf{1}_{t\geq T_3}(t)\left[u_{2,j} + \hat{z}_{2,2,j} + \lambda_1 \lceil \hat{V}_{2,j} - \hat{z}_{2,1,j} \rfloor^{1/2}\right], \\ \dot{\hat{z}}_{2,2,j} \in \mathbf{1}_{t\geq T_3}(t)\,\lambda_2\,\mathrm{sign}(\hat{V}_{2,j} - \hat{z}_{2,1,j}), \end{cases} \qquad j \in [\![1, n_2]\!], \tag{10.2g}$$

$$\hat{\Phi}_2 := \begin{bmatrix} S(y) \\ S(\hat{V}_2) \\ 1 \end{bmatrix}, \qquad [\hat{W}_{21} \mid \hat{W}_{22} \mid \hat{h}_2] := \left(\int_{T_4}^{T_5} (\hat{z}_{2,2} - A_2\hat{V}_2)\hat{\Phi}_2^\top\right)\left(\int_{T_4}^{T_5} \hat{\Phi}_2\hat{\Phi}_2^\top\right)^{-1}. \tag{10.2h}$$

Theorem 10.1. *Under assumptions (H1)–(H4), the identification protocol (10.2) recovers all unknown parameters of (10.1) exactly. Specifically, for any $\rho > 0$ and any time sequence $0 < T_0 < T_0' < T_1 < T_2 < T_3 < T_4 < T_5$ with sufficiently large interval lengths (depending only on ρ and the known data $(\bar{A}_1, \bar{h}_1, \bar{h}_2, \bar{W}, A_2, |\underline{R}|, |\bar{R}|)$), there exist observer gains (λ_1, λ_2), $\tau > 0$ and piecewise bounded controls (u_1, u_2) such that for every solution of (10.1) initialized in $B(0, \rho)$:*

1. *$[\hat{A}_1 \mid \hat{h}_1] = [A_1 \mid h_1]$, where $[\hat{A}_1 \mid \hat{h}_1]$ is defined by (10.2b);*
2. *$\hat{W}_{11} = W_{11}$, where $\hat{W}_{11}$ is defined by (10.2c);*
3. *$\hat{W}_{12} = W_{12}$, where $\hat{W}_{12}$ is defined by (10.2e);*
4. *$\hat{V}_2(t) = V_2(t)$ for all $t \geq T_3$, where $\hat{V}_2$ is defined by (10.2f);*
5. *$[\hat{W}_{21} \mid \hat{W}_{22} \mid \hat{h}_2] = [W_{21} \mid W_{22} \mid h_2]$, where $[\hat{W}_{21} \mid \hat{W}_{22} \mid \hat{h}_2]$ is defined by (10.2h).*

Remark 10.2. *The identification problem is well-posed only because S is nonlinear. For example should S be a linear function, the first equation of (10.1) would read $\dot{V}_1 = (A_1 + W_{11})V_1 + W_{12}V_2 + u_1 + h_1$, and the parameters A_1 and W_{11} would be structurally unidentifiable.*

Remark 10.3. *The strict monotonicity of S on $[\underline{R}, \bar{R}]$ (hence the existence of S^{-1}) is not required for items 1–3 of Theorem 10.1. The estimators (10.2b)–(10.2e) depend only on $S(V_1)$ and $S(V_2)$, which take prescribed values by construction and do not require inversion. Invertibility of S is first needed in item 4, where $\hat{V}_2 = S^{-1}(\hat{W}_{12}^{\dagger}\phi)$ (10.2f), and consequently in item 5.*

Remark 10.4. *The batch estimators (10.2b)–(10.2h) admit equivalent online formulations via recursive least squares: the integral $\left(\int \Phi\Phi^\top\right)^{-1}\int \hat{z}\,\Phi^\top$ is replaced by the standard RLS update running on the same sliding-mode signal $\hat{z}$. The batch form is retained here because it makes the identification argument transparent.*

10.3 TECHNICAL LEMMAS

The four lemmas below are direct consequences of the results of Chapter 8, applied to the two populations of (10.1). Throughout this section $\rho > 0$ bounds the initial conditions.

Lemma 10.5 (Disconnection of V_2). *Let $\rho, T_0 > 0$. There exists $\mu^\star > 0$, depending only on $(A_2, \bar{W}, \bar{h}_2, \rho, T_0)$, such that for any $\mu \geq \mu^\star$, the constant control*

$$u_2(t) = -\mu\,\mathbf{1}_{n_2},$$

for any $u_1 \in L^\infty$ and any solution of (10.1) initialized in $B(0, \rho)$, yields

$$S(V_2(t)) = 0 \quad \forall t \geq T_0.$$

Moreover, $\|V_2(t)\|_\infty \leq R_\mu$ for all $t \geq 0$, where R_μ is the uniform bound of (8.25).

Proof. Since A_2 is known Hurwitz and $B_2 = I_{n_2}$, the second equation of (10.1) has the structure of (8.2) with $B = I_{n_2}$ and perturbation $d := W_{21}S(V_1) + W_{22}S(V_2) + h_2$ bounded by $b_d = 2\bar{W} + \bar{h}_2$. The negative orthant is reachable by $-\mu\mathbf{1}_{n_2}$ since $B = I_{n_2}$. The result follows from Lemma 8.14 with $\mathcal{E} = 0$. □

Lemma 10.6 (Steering V_1 into the linear regime and PE generation)**.** *Let $\rho, T_0 > 0$ and $\gamma^* \in C^{1,L}(\mathbb{R}, (-\infty, \underline{R})^{n_1})$ persistently exciting. There exist $\mu_1^\star, \mu_2^\star > 0$, depending only on $(\bar{A}_1, \bar{W}, \bar{h}_1, \bar{h}_2, A_2, \|\gamma^*\|_{C^{1,L}}, \rho, T_0)$ such that for any $\mu_1 \geq \mu_1^\star$ and $\mu_2 \geq \mu_2^\star$, the control*

$$u_1(t) = -\mu_1\big(y(t) - \gamma^*(t)\big) + \dot{\gamma}^*(t), \qquad u_2(t) = -\mu_2\,\mathbf{1}_{n_2},$$

yields, for any solution of (10.1) initialized in $B(0, \rho)$:

- *(i) $S(V_2(t)) = 0$ for all $t \geq T_0$, and $\|V_1(t) - \gamma^*(t)\|_\infty \leq C/(\mu_1 - \bar{A}_1)$ for all $t \geq T_0$ after a transient of duration $O(1/(\mu_1 - \bar{A}_1))$, where C depends only on $(\bar{A}_1, \bar{W}, \bar{h}_1, \|\gamma^*\|_{C^{1,L}})$; in particular both the bound and the transient duration can be made as small as desired by taking μ_1 large; since $\gamma^*(t) < \underline{R}$ strictly, choosing μ_1 large enough that $C/(\mu_1 - \bar{A}_1) < \inf_t(\underline{R} - \|\gamma^*(t)\|_\infty)$ ensures $S(V_1(t)) = 0$ for all $t \geq T_0$;*
- *(ii) on $[T_0, T_0']$ with $T_0' - T_0$ large enough, Φ_1 satisfies the PE condition and the estimator (10.2b) satisfies $[\hat{A}_1 \mid \hat{h}_1] = [A_1 \mid h_1]$.*

Proof. (i): Disconnection of V_2 follows from Lemma 10.5. For V_1: let $e = V_1 - \gamma^*$. Once $S(V_2) = 0$, the tracking error satisfies

$$\dot{e} = (A_1 - \mu_1 I)e + A_1\gamma^* + W_{11}S(V_1) + h_1,$$

where $\|A_1\gamma^* + W_{11}S(V_1) + h_1\| \leq \bar{A}_1\|\gamma^*\|_\infty + \bar{W}\sqrt{n_1} + \bar{h}_1 =: C$ (since $\|S(V_1)\|_2 \leq \sqrt{n_1}$, each component of $S(V_1)$ lying in $[0, 1]$). Since $\|A_1\| \leq \bar{A}_1$, the inequality $e^\top(A_1 - \mu_1 I)e \leq (\bar{A}_1 - \mu_1)\|e\|^2$ holds, so for $\mu_1 > \bar{A}_1$ this gives $\|e(t)\| \leq C/(\mu_1 - \bar{A}_1)$. Since $\gamma^*(t) < \underline{R}$ strictly, choosing μ_1 large enough ensures $V_1 = \gamma^* + e < \underline{R}$ component-wise, hence $S(V_1(t)) = 0$ for all $t \geq T_0$.

(ii): On the window where $S(V_1) = S(V_2) = 0$, the sliding mode gives $\hat{z}_{1,2} = \dot{V}_1 - u_1 = A_1V_1 + h_1 = [A_1 \mid h_1]\Phi_1$ exactly. Since $V_1 = \gamma^* + O(1/\mu_1)$, the regressor $\Phi_1 = [\gamma^*; 1] + O(1/\mu_1)$. For μ_1 large enough, Φ_1 satisfies the PE condition on $[T_0, T_0']$ (since $[\gamma^*; 1]$ is PE by assumption and PE is open in L^∞), and the estimator (10.2b) is exact. □

Lemma 10.7 (Cycling saturation values)**.** *Let $\rho, T > 0$, $\tau > 0$, and $t_1 < \cdots < t_m \subset [0, T]$ with $t_{i+1} - t_i > \tau$.*

- *(i) With A_1, h_1 known and $u_2 = -\mu_2\mathbf{1}_{n_2}$ (so $S(V_2) = 0$), there exists a piecewise constant control u_1 such that for any solution initialized in $B(0, \rho)$,*

$$S(V_1(t)) = e_i \quad \forall t \in [t_i, t_i + \tau], \quad i \in [\![1, n_1]\!].$$

(ii) *With A_1, h_1, W_{11} known and u_1 keeping $S(V_1) = 0$, there exists a piecewise constant control u_2 such that for any solution initialized in $B(0, \rho)$,*

$$S(V_2(t)) = e_j \quad \forall\, t \in [t_j, t_j + \tau], \quad j \in [\![1, n_2]\!].$$

Proof. Both parts are direct applications of Lemma 8.17 of Chapter 8. In (i), the first equation with $A = A_1$ known and $B = I_{n_1}$ satisfies the structural hypotheses, and $e_i \in \mathcal{A}$ since $B = I_{n_1}$. In (ii), the second equation with $A = A_2$ known and $B = I_{n_2}$ likewise satisfies them. In both cases the coupling terms are treated as bounded perturbations of magnitude at most $2\bar{W} + \bar{h}_i$. □

Lemma 10.8 (Tracking in the invertible region and second-layer PE)**.** *Assume A_1, h_1, W_{11}, W_{12} all known. Let $\delta \in (0, 1/2)$, $\rho > 0$. Let $\gamma_1 \in C^{1,L}(\mathbb{R}, \Omega^{n_1}_{S,\delta})$ and $\gamma_2 \in C^{1,L}(\mathbb{R}, \Omega^{n_2}_{S,\delta})$ be references such that $[S(\gamma_1); S(\gamma_2); 1]$ is persistently exciting. There exist $\mu_1^\star, \mu_2^\star > 0$, depending only on $(\bar{A}_1, \bar{W}, \bar{h}_1, \bar{h}_2, A_2, \|\gamma_1\|_{C^{1,L}}, \|\gamma_2\|_{C^{1,L}}, \rho, \delta)$, such that for any $\mu_1 \geq \mu_1^\star$ and $\mu_2 \geq \mu_2^\star$, the controls*

$$u_1(t) = -\mu_1\big(y(t) - \gamma_1(t)\big) + \dot{\gamma}_1(t), \qquad u_2(t) = \eta\big(W_{12}^\dagger \phi(t);\, \mu_2,\, \gamma_2(t),\, \delta\big), \tag{10.3}$$

where ϕ is defined by (10.2d) and η is the feedback law (8.35) of Chapter 8, yield for any solution initialized in $B(0, \rho)$:

(i) *there exists a finite time T_3, depending only on the same parameters as $\mu_1^\star, \mu_2^\star$, such that $V_2(t) \in \Omega^{n_2}_{S,\delta}$ and $\hat{V}_2(t) = V_2(t)$ for all $t \geq T_3$;*

(ii) *$\|V_1(t) - \gamma_1(t)\|_\infty \leq C_1/(\mu_1 - \bar{A}_1)$ and $\|V_2(t) - \gamma_2(t)\|_\infty \leq C_2/\mu_2$ for all t after a transient of duration $O(1/(\mu_i - \bar{A}_i))$, where C_1, C_2 depend only on $\bar{A}_1, \bar{W}, \bar{h}_1, \bar{h}_2, \|\gamma_1\|_{C^{1,L}}, \|\gamma_2\|_{C^{1,L}}$; in particular, by taking μ_1, μ_2 large, both bounds and the transient duration can all be made as small as desired;*

(iii) *for any $T_4 \geq T_3$ and any $T_5 > T_4$ with $T_5 - T_4$ sufficiently large (depending on the same parameters), $\hat{\Phi}_2$ satisfies the PE condition on $[T_4, T_5]$ and the estimator (10.2h) is exact.*

Proof. Since W_{12} has rank n_2, the signal $S(V_2) = W_{12}^\dagger \phi$ is directly observable from (10.2d). Lemma 8.18 of Chapter 8, applied to the second equation with $\tilde{y} = W_{12} S(V_2) = \phi$ as the observed output, provides the control η and establishes (i). Once $V_2 \in \Omega^{n_2}_{S,\delta}$, S is invertible and $\hat{V}_2 = S^{-1}(W_{12}^\dagger \phi) = V_2$ exactly. For the V_1 part of (ii): since u_1 enters directly, the feedback $u_1 = -\mu_1(y - \gamma_1) + \dot{\gamma}_1$ makes $(A_1 - \mu_1 I)$ satisfy $e^\top (A_1 - \mu_1 I) e \leq (\bar{A}_1 - \mu_1)\|e\|^2$ for any $\mu_1 > \bar{A}_1$. Since the disturbance $A_1\gamma_1 + W_{11}S(V_1) + W_{12}S(V_2) + h_1$ is uniformly bounded by $C_1 := \bar{A}_1 \|\gamma_1\|_\infty + \bar{W}(\sqrt{n_1} + \sqrt{n_2}) + \bar{h}_1$ (using $\|S(V_i)\|_2 \leq \sqrt{n_i}$ since each component of $S(V_i)$ lies in $[0, 1]$), the error $e_1 = V_1 - \gamma_1$ satisfies $\dot{e}_1 \leq -(\mu_1 - \bar{A}_1)\|e_1\| + C_1$, which by the comparison lemma gives $\|e_1(t)\| \leq \|e_1(0)\| e^{-(\mu_1 - \bar{A}_1)t} + C_1/(\mu_1 - \bar{A}_1)$. Both the bound $C_1/(\mu_1 - \bar{A}_1)$ and the transient duration are $O(1/\mu_1)$ and can be made as small as desired by choosing μ_1 large enough, establishing the claim of (ii). The bound for V_2 follows by the same argument applied to the η law of Lemma 8.18.

For (iii): the tracking in (ii) gives $V_1 = \gamma_1 + O(1/\mu_1)$ and $V_2 = \gamma_2 + O(1/\mu_2)$, so $\hat{\Phi}_2 = [S(\gamma_1); S(\gamma_2); 1] + O(1/\mu)$ in L^∞. Since $[S(\gamma_1); S(\gamma_2); 1]$ is PE by assumption and the PE condition is open in L^∞, $\hat{\Phi}_2$ satisfies the PE condition on any sufficiently long window $[T_4, T_5]$ with $T_4 \geq T_3$, for μ_1, μ_2 large enough. □

10.4 PROOF OF THEOREM 10.1

Let $\rho > 0$ bound the initial conditions.

Step 1: Silencing V_2. For the fixed T_0 from the time sequence, apply Lemma 10.5 to obtain $\mu_2^\star$ and hence the control $u_2 = -\mu_2 \mathbf{1}_{n_2}$ (any $\mu_2 \geq \mu_2^\star$) such that $S(V_2(t)) = 0$ for all $t \geq T_0$. The state bound R_μ from that lemma determines the uniform bound on $\ddot{V}_1 - \dot{u}_1$; for (λ_1, λ_2) chosen large enough relative to this bound, observer (10.2a) gives $\hat{z}_{1,2}(t) = \dot{V}_1(t) - u_1(t)$ for all $t \geq T_0$.

Step 2: Recovering A_1 *and* h_1. Apply Lemma 10.6 with a PE reference $\gamma^* \in (-\infty, \underline{R})^{n_1}$. The feedback control $u_1 = -\mu_1(y - \gamma^*) + \dot{\gamma}^*$ forces $V_1 = \gamma^* + O(1/\mu_1)$ uniformly from T_0, so $S(V_1) = 0$ and Φ_1 is PE on $[T_0, T_0']$; the estimator (10.2b) is exact.

Step 3: Recovering W_{11}. With $\hat{A}_1, \hat{h}_1$ known, Lemma 10.7(i) is applied on disjoint windows $[s_i, s_i + \tau] \subset [T_0', T_1]$ to force $S(V_1) = e_i$, while $u_2 = -\mu_2 \mathbf{1}_{n_2}$ keeps $S(V_2) = 0$. The residual $\phi_1 = \hat{z}_{1,2} - \hat{A}_1 y - \hat{h}_1 = W_{11} S(V_1)$ on those windows makes the estimator (10.2c) exact.

Step 4: Recovering W_{12}. With A_1, h_1, W_{11} known, Lemma 10.7(ii) provides a piecewise constant u_2 cycling $S(V_2)$ through $e_1, \ldots, e_{n_2}$ on windows $[t_j, t_j + \tau] \subset [T_1, T_2]$, with u_1 keeping $S(V_1) = 0$. The residual $\phi = W_{12} S(V_2)$ makes the estimator (10.2e) exact.

Step 5: Recovering V_2 *and the second-layer parameters.* Lemma 10.8(i) provides u_2 on $[T_2, T_3]$ bringing V_2 into $\Omega_{S,\delta}^{n_2}$; from T_3 onward $\hat{V}_2 = V_2$ exactly by (10.2f). Observer (10.2g) is activated at T_3; the same gains (λ_1, λ_2) suffice. On $[T_4, T_5]$, the two tracking feedbacks $u_1 = -\mu_1(y - \gamma_1) + \dot{\gamma}_1$ and $u_2 = \eta(W_{12}^\dagger \phi; \mu_2, \gamma_2, \delta)$ drive (V_1, V_2) toward $\gamma_1 \in \Omega_{S,\delta}^{n_1}$ and $\gamma_2 \in \Omega_{S,\delta}^{n_2}$ respectively, where the pair is chosen so that $[S(\gamma_1); S(\gamma_2); 1]$ is PE. Lemma 10.8(iii) then ensures $\hat{\Phi}_2$ satisfies the PE condition, and the identity $\hat{z}_{2,2} - A_2 \hat{V}_2 = [W_{21} \mid W_{22} \mid h_2] \hat{\Phi}_2$ makes the estimator (10.2h) exact.

Sliding-mode gains. At each phase the bound on $\ddot{V}_1 - \dot{u}_1$ (resp. $\ddot{V}_2 - \dot{u}_2$ after T_3) is finite and determined by known bounds, so (λ_1, λ_2) can always be chosen large enough to maintain the finite time convergence at each phase.

10.5 NUMERICAL SIMULATION

The identification procedure of Theorem 10.1 is illustrated on a randomly generated instance of (10.1) with $n_1 = 3$ and $n_2 = 2$. The matrices $A_1 = -D_1$ and $A_2 = -D_2$ are diagonal Hurwitz, with diagonal entries drawn uniformly from $[2, 5]$ and $[1, 2]$. The entries of W_{11}, W_{21}, W_{22}, and $W_{12} \in \mathbb{R}^{3\times 2}$ are drawn i.i.d. from $\mathcal{N}(0, 1)$, with W_{12} resampled until it has rank 2. The biases h_1, h_2 are zero, and S is a smooth sigmoidal saturation with known levels $\underline{R} < \bar{R}$.

The simulation runs over $[0, T_f] = [0, 80]$ with step $\Delta t = 5 \times 10^{-3}$, following the timeline $T_0 < T_0' < T_1 < T_2 < T_3 < T_4 < T_5$ of Theorem 10.1. Since A_1 is known and $h_1 = 0$ in this instance, Step 2 of the protocol (identification of $[A_1 \mid h_1]$) is skipped; accordingly $T_0' = T_0 = 8$, and the remaining breakpoints are $T_1 = 24$, $T_2 = 56$, $T_5 = 64$.

- $[0, T_0] = [0, 8]$: $u_2 = -\mu \mathbf{1}_{n_2}$ with $\mu = 4$ drives $S(V_2)$ to zero; observer (10.2a) runs with gains $\lambda_1 = 30$, $\lambda_2 = 37$.

- $[T_0', T_1] = [8, 24]$: $S(V_1)$ is cycled through e_1, e_2, e_3 by piecewise constant controls; estimator (10.2c) recovers W_{11}.
- $[T_1, T_2] = [24, 56]$: $S(V_2)$ is cycled through e_1, e_2 over $n_2 = 2$ sub-intervals; each column of W_{12} is recovered independently: on the window where $S(V_2) = e_j$, the residual $\phi = W_{12}e_j$ is the j-th column, giving $\hat{W}_{12}[\,\cdot\,, j] := \frac{1}{\tau}\int_{t_j}^{t_j+\tau} \phi \, \mathrm{d}t$.
- $[T_2, T_5] = [56, 64]$: V_2 is steered into $\Omega^{n_2}_{S,\delta}$. The time $T_3 \in (56, 64)$ at which V_2 first enters this region is determined by the dynamics and is not fixed in advance; once reached, $\hat{V}_2 = S^{-1}(\hat{W}_{12}^{\dagger}\phi)$ is exact. Observer (10.2g) is activated at T_3 and reaches its sliding surface at time $T_4 \in (T_3, 64)$, also determined dynamically. Estimator (10.2h) recovers W_{21}, W_{22}, h_2 on $[T_4, T_5]$.
- $[T_5, T_f] = [64, 80]$: validation (closed-loop tracking error monitored).

Remark 10.9. *The simulation uses the running-average form of the estimators; as noted in Remark 10.4, this is algebraically equivalent to the batch formulas of Theorem 10.1.*

Figure 10.1 shows the errors $\|\hat{W}_{11} - W_{11}\|$, $\|\hat{W}_{12} - W_{12}\|$, and $\|\hat{V}_2 - V_2\|$. Each reaches zero at the end of its identification window.

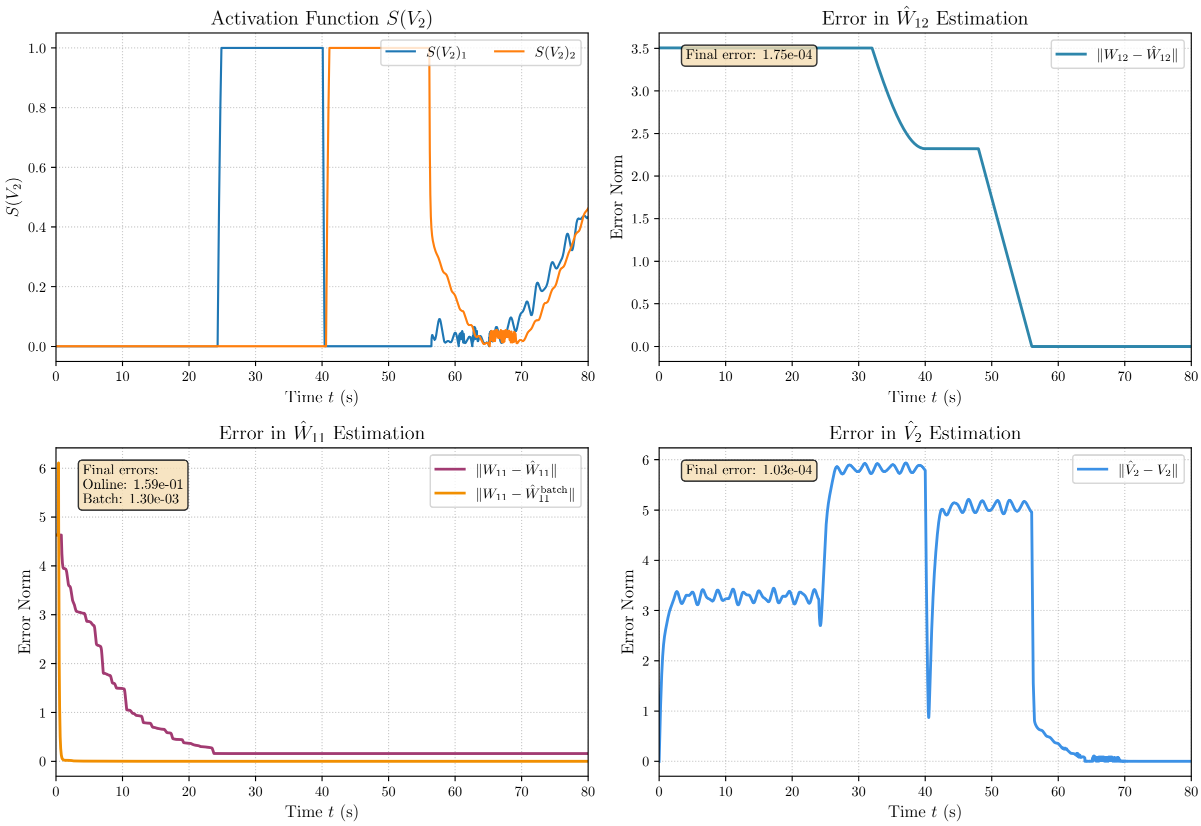


Figure 10.1: Estimation errors for the simulation instance. Since A_1 is known and $h_1 = 0$, Stage 2 is skipped. From top to bottom: errors $\|\hat{W}_{11} - W_{11}\|$, $\|\hat{W}_{12} - W_{12}\|$, and $\|\hat{V}_2 - V_2\|$. Each error reaches zero at the end of its identification window.

11

IDENTIFICATION OF HIERARCHICAL SATURATED NETWORKED SYSTEMS (PUBLISHED CDC 2025)

This chapter reproduces the article [9], published in *IEEE Conference on Decision and Control (CDC) 2025*.

We recall here that although this chapter comes last, the results it contains were obtained before the more general results of the previous chapters, and that the latter were developed in an attempt to extend and refine the method proposed in this article. It is therefore the most "crude" version of the results.

We address the challenge of parameter identification in networks of neural mass models. We consider a system of n interconnected populations, structured in successive layers, where measurements are available from only one. Traditional offline methods struggle with parameter ambiguity and time-dependent variability. We propose an online approach that exploits the system's nonlinear characteristics, particularly saturation functions, to enable effective parameter recovery. By designing specific control inputs applied to targeted nodes, it becomes possible to isolate and retrieve parameters effectively. Crucially, we show that this process must be conducted online to achieve full identification, as control inputs must adapt to previously unknown states and parameters.

11.1 INTRODUCTION

The application of control theory techniques to neuroscience presents a rich and challenging problem. Its diverse applications range from brain health monitoring [149] to brain-computer interfacing [150]. For example, feedback control has been employed to stabilize brain signals and mitigate Parkinson's disease symptoms [62, 47].

These methods and techniques are subject to the choice of a relevant model. We focus here on neural mass models, particularly those based on the Wilson–Cowan framework [168]. They serve as powerful tools for studying the macroscopic dynamics of neuronal populations [61, 168]. See [36] for review on large scale models of the brain.

All dynamical models depend on a certain number of parameters whose values influence the behavior and the response to controls. It is therefore critical to know, exactly or approximately, the value of these paramaters. Some of them may be known from physical considerations, but most of the parameters have to be estimated from indirect measurement via *parameter identification*. By identification we mean any method or process that uses measurements of various outputs of the system corresponding to some inputs (or excitation) to deduce some estimate of the parameters. The choice of the excitation, or experiment design, may be an important part of the process.

Parameter identification in neural mass models is a challenging problem due to the non-linear nature of the system. Offline methods struggle with issues such as parameter ambiguity, where different parameter sets can produce equally good representations of the data [20]. Moreover, this approach fails to account for time-dependent variations in physiological parameters [150]. This underscores the necessity of developing tools that leverage online approaches.

Since we are dealing with a case where both the state is not fully measured and parameters are unknown, it is natural to resort to adaptive observers, see *e.g.* [69]. This is done for instance in [40, 43], even for a distributed state in the first case. This was done under various sorts of contraction and persistent excitation assumptions, that are not easy to check in general, and are very weakly satisfied in realistic neural systems where saturation effects induce non-monotonic behaviors, leading to disconnection between variables.

We want to propose an identification method that includes experiment design, i.e. proposes explicit control that facilitates identification; see [17] for a discussion of this problem and [131, 164] for experiment design for systems with interconnections in a different context from ours.

In this work, we study a network of n interconnected neural populations, structured in successive layers, governed by Wilson-Cowan-like equations, where measurements are available from only one population. We propose an identification approach that exploits the non-linear characteristics of the system, particularly the role of saturation functions that are common in these models [36, 168], to establish a one-to-one correspondence between the output of the system and its parameters. By designing specific control inputs applied to selected nodes, we ensure that all parameters can be retrieved effectively. Crucially, we show that the parameter estimation must be done online, as it is inherently recursive. Indeed, control inputs must account for previously unknown states or parameters to achieve full system identification. This highlights the necessity of online estimation techniques capable of adapting to unknown and time-varying dynamics. Our approach addresses key limitations of offline methods and introduces new possibilities for real-time neural modulation and feedback control.

NOTATIONS 1. For any two integers $p < q$, we denote by $[\![p, q]\!]$ the set of integers $[\![p, q]\!] = \{i \in \mathbb{Z},\ p \leq i \leq q\}$.
2. Given an element V in $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_p} = \mathbb{R}^{n_1 + \cdots + n_p}$ and q in $[\![1, p-1]\!]$, we let $V^q = (V_1, \ldots, V_q)$.
3. Given a function $g : \mathbb{R} \to \mathbb{R}$ and a vector $z \in \mathbb{R}^n$, we denote by $g(z) \in \mathbb{R}^n$ the vector obtained by applying g element-wise, i.e., $g(z) = (g(z_1), \ldots, g(z_n))$.

11.2 MODEL DISCUSSION

The i^{th} population, $1 \leq i \leq p$, is described by the state variable $V_i \in \mathbb{R}^{n_i}$; the state of the network is $V = (V_1, \dots, V_p) \in \mathbb{R}^N = \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_p}$ with $n_1 \geq \cdots \geq n_p$, $N_k = n_1 + \cdots + n_k$, $N = N_p$.

The dynamics are given by

$$\begin{cases} \dot{V} = f(V, u, \Theta), \\ y = V_1, \end{cases} \tag{11.1}$$

where the output $y = V_1$ indicates that we entirely measure the first population, and this is the only measurement. The right-hand side $f = (f_1, \dots, f_p)$ has more structure: each f_i, $1 \leq i \leq p$ is given by

$$f_i(V, u, \Theta) = -V_i + \sum_{j=1}^{\min\{p, i+1\}} W_{ij} S(V_j) \; + \; u_i \,. \tag{11.2}$$

Assumption 11.1. *The function $S : \mathbb{R} \to \mathbb{R}$ is a known Lipschitz-continuous, piecewise C^1 function satisfying*

$$x \leq \underline{R} \Rightarrow S(x) = 0 \quad \textit{and} \quad x \geq \bar{R} \Rightarrow S(x) = 1 \tag{11.3}$$

for known real numbers $\underline{R} < \bar{R}$, and increasing in $[\underline{R}, \bar{R}]$.

The function S is modeled by a modified sigmoid-type function which attains its limits. A possible candidate will be the clipped ReLU function $S(\lambda) = (\lambda - \underline{R})/(\bar{R} - \underline{R})$, $\lambda \in [\underline{R}, \bar{R}]$). This assumption is often made in the literature. (see, e.g, [54]).

Define, for convenience,

$$\begin{aligned} &\Gamma_k := \{(i,j) \in [\![1,k]\!]^2, \, 1 \leq j \leq \min(i+1, k)\}, \\ &\text{and } \; \Theta_k := \big(W_{ij}\big)_{(i,j) \in \Gamma_k} \,, \; 1 \leq k \leq p \,, \end{aligned} \tag{11.4}$$

and let Q_k be the total number of entries in Θ_k; one may consider that Θ_k is an element of $\mathbb{R}^{Q_k}$. The parameters of the vector field f are the matrices $W_{i,j} \in \mathbb{R}^{n_i \times n_j}$, that can be lumped into $\Theta = \Theta_p \in \mathbb{R}^{Q_p}$. They are unknown but we assume the following.

Assumption 11.2. *For some known number $\bar{W} > 0$, one has*

$$\|W_{i,j}\|_\infty \leq \bar{W}, \quad (i,j) \in \Gamma_p \,, \tag{11.5}$$

$$W_{k,k+1} \textit{ is left-invertible}, \; 1 \leq k \leq p-1 \,. \tag{11.6}$$

We call a system of type (11.1)-(11.2) a *p-layer system*, and we will sometimes call its *q-layer subsystem* ($q \leq p$) the system with state $(V_1, \dots, V_q)$ obtained by discarding $V_{q+1}, \dots, V_p$ and removing the term $W_{q,q+1} S(V_{q+1})$ in $\dot{V}_q$.

In the rest of the article we will consider controls that are continuous by part. From Cauchy-Lipschitz theorem, solutions of (11.1) are well defined, global in time. Moreover, if u_i are bounded, due to the contraction induced by the $-x$ term and the boundedness of the sigmoid S, there exist an invariant attracting bounded set $\mathcal{V}$ for the dynamic $f(v, u)$ which denotes the right hand side of (11.1). We assume from hereafter that all solutions of (11.1) are bounded and belong to the set $\mathcal{V}$.

11.3 MAIN RESULTS

In this section we present the main contribution of this article in the form of the following theorem

Theorem 11.3. *Assume hypotheses 11.1-11.2 hold. Let $T > 0, \rho > 0$. For $k \in \{1, \dots, p\}$ there exists*

- *a family of maps $U_t^k : \mathbb{R}^{Q_{k-1}} \times C([0,t]; \mathbb{R}^{n_1}) \to \mathbb{R}^{N_k}$,*
- *a map $\Phi_t^k : \mathbb{R}^{Q_{k-1}} \times C([0,t], \mathbb{R}^{n_1}) \to \mathbb{R}^{Q_k} \times C([0,t], \mathbb{R}^{N_k})$,*

such that for any solutions of (11.1) with initial conditions in $B_N(0,\rho)$ and output feedback control

$$u(t) = U_t^k(\Theta_{k-1}, y|_{[0,t]}), \tag{11.7}$$

we have that

$$\phi_t^k(\Theta_{k-1}, y|_{[0,t]}) = \Big(\Theta_k, V^k(t)\Big), \quad \forall\, t \geq T. \tag{11.8}$$

In particular, there exist a map $\Psi : C([0,kT], \mathbb{R}^{n_1}) \to \mathbb{R}^Q$ such that

$$\Psi(y|_{[0,kT]}) = \Theta. \tag{11.9}$$

The above theorem describe that from the knowledge of the output V_1 and the parameters Θ_k, there exists a control for which the measurement is rich enough to let you obtain Θ_j, for any degree of depth desired $j \geq k+1$. The specific statement of the theorem formalizes the idea that, through a sequence of finite open-loop experiments, it is possible to recover unknown parts of the parameter set. Once the parts of the parameters are established, a new control can then be designed to capture the remaining unknown parameters.

11.3.1 *Structure of the proof*

The remainder of this paper is structured to progressively develop the tools required to prove the main theorem. In *Section 11.4* we analyze the control of the nonlinearities in (11.1) through an open-loop strategy. We expend on this to show that we can "disconnect" lower layers from the network. We also define output feedback controls for identifying the terms W_{kj} at each layer $k \in [\![1,p]\!]$ for $j \in [\![1,k]\!]$. In *Section 11.5* we identify the interaction terms $W_{i,i+1}$ for $i \in [\![1,p-1]\!]$. In *Section 11.6* we provide the tools to recover the state of successive layers up to order k using measurements from the first layer $y = V_1$, when the parameters Θ_k are known up to a certain order k. Finally in *Section 11.7* we combine the results to prove Theorem 11.3.

We numerically illustrate the method in Section 11.8, and recall some results on finite time observers in Appendix 11.9.

11.4 1-LAYER SYSTEM

In this section, we focus on the following 1-layer system

$$\dot{x} = -x + \psi(x,t) + u, \tag{11.10}$$

where $x \in \mathbb{R}^n$, $\psi \in C_b(\mathbb{R}^n \times \mathbb{R}, \mathbb{R}^n)$ and $u \in C_b(\mathbb{R}, \mathbb{R}^n)$. Solutions of (7.40) are well defined and global in time. Indeed, due to the contraction induced by the $-x$ term, we have for any $R \geq R_0 := \|\psi\|_\infty + \|u\|_\infty > 0$ the ball $B_n(0, R)$ is an invariant set for (7.40).

11.4.1 *Weak controllability of the non-linearities*

In this section, we prove that we can choose constant controls steering $S(x)$ towards any vector whose components are either 0 or 1. This can be achieved in finite time and without *a priori* knowledge of x.

Lemma 11.4. *Let $T > 0$, $\rho > 0$, $m > 0$, and $w \in \mathcal{A}_n := \{0, 1\}^n \subset \mathbb{R}^n$. Consider the constant control u^w*

$$u_j^w := \begin{cases} \mu, & \text{if } w_k = 1, \\ -\mu, & \text{if } w_k = 0, \end{cases} \tag{11.11}$$

with $\mu \geq \mu^*$, where

$$\mu^* := m + \frac{\rho + e^T \max\{|\underline{R}|, |\bar{R}|\}}{e^T - 1}. \tag{11.12}$$

Then, for any ψ such that $\|\psi\|_\infty \leq m$, solutions of

$$\dot{x} = -x + \psi(x, t) + u^w. \tag{11.13}$$

initialized in $B_n(0, \rho)$, satisfy $S(x(t)) = w$ when $t \geq T$. Here, $\bar{R}$ and $\underline{R}$ are set in (11.3).

Proof. Since the components of x are decoupled, without loss of generality, we can assume $n = 1$. Let us show that if $w = 1$, then $x(t) \geq \bar{R}$ for all times $t \geq T$. From the expression of μ^* in (11.12), we have $\dot{x} \geq -x + (\rho + \bar{R})/(e^T - 1)$. The claim follows by Grönwall's lemma and the fact that $x(0) \geq -\rho$. A similar argument shows that $x(t) \leq \underline{R}$ if $w = 0$, completing the proof. □

11.4.2 *Cutting Layers*

Denote the restriction $f^q = (f_1, \dots, f_q)$. Recall that for all $i \in [\![1, p]\!]$, $f_i(V, u, \Theta) = f_i(V^{i+1}, u_i, \Theta_{i+1})$.

From the above lemma, it follows that we can reduce p-layer systems to a q-layer system, $q < p$, without any knowledge of the parameters or the states of the removed layers. This is encapsulated in the following.

Corollary 11.5. *Let $T > 0$, $\rho > 0$, and $q \in [\![1, p-1]\!]$. Then, there exists constant controls $u_{q+1}, \dots, u_p$ such that for any solution V of (11.1) with initial condition in $B_{n_1}(0, \rho) \times \cdots \times B_{n_p}(0, \rho)$, the vector V^q follows for $t > T$ the dynamics of a q-layer system (see (11.2)) with parameter Θ_q. That is, in the dynamic f_q, for $t \geq T$,*

$$f_q(V(t), u(t), \Theta) = -V_q(t) + \sum_{i=1}^{q} WS(V_i(t)) + u_q(t).$$

11.4.3 *Identification of the nonlinearities under structural assumptions*

In this section, we show how to identify the non-linearity ψ in the particular case where $\psi(x,t) = WS(x)$, for $W \in \mathbb{R}^{n\times n}$. We suppose here that we can estimate $S(x)$, it will be treated in the next section. We consider system (11.10) with output

$$\begin{cases} \dot{x} = -x + WS(x) + u, \\ y = S(x). \end{cases} \tag{11.14}$$

To estimate the parameter W, we first construct a class of output feedback controls u^*. We show that under this class of functions, solutions of (11.14) initialized in a known bounded domain remain in the invertible region of S. Moreover, we demonstrate that this class of controls can make $S(x)$ a persistently exciting function, allowing the estimation of W in finite time.

Given $L > 0$, denote by $C^{1,L}$ the set of those C^1 trajectories whose derivative is bounded by L. Given $0 < \delta < 1/6$, $l > 0$, and $\gamma \in C^{1,L}\left(\mathbb{R}, (\underline{R} + \delta, \bar{R} - \delta)^n\right)$, we then define the following piecewise affine function

$$\eta(\sigma; l, \delta, \gamma) = \begin{cases} l & \text{if } \sigma \leq \ \delta, \\ \alpha_1(\sigma; l, \delta) & \text{if } \delta \leq \sigma \leq 2\delta, \\ -l\left(S^{-1}(\sigma) - \gamma\right) & \text{if } 2\delta \leq \sigma \leq 1 - 2\delta, \\ \alpha_2(\sigma; l, \delta) & \text{if } 1 - 2\delta \leq \sigma \leq 1 - \delta, \\ -l & \text{if } \sigma \geq 1 - \delta. \end{cases}$$

Here, α_1 and α_2 are affine functions ensuring continuity of η, and $S^{-1} : [0,1] \to [\underline{R}, \bar{R}]$ is the inverse of S in the invertible region. Define now the output feedback control

$$u^*(y) = \eta(y; l, \delta, \gamma). \tag{11.15}$$

Define the following set

$$\Omega_\delta := \left(S^{-1}(3\delta), S^{-1}(1 - 3\delta)\right). \tag{11.16}$$

Lemma 11.6. *Let $T, \rho, \bar{W}, L > 0$ and $0 < \delta < 1/6$. For any $\varepsilon > 0$ and any $\gamma \in C^{1,L}\left(\mathbb{R}, \Omega_\delta^n\right)$, there exist $l^* > 0$ such that for any $l > l^*$, and any $W \in \mathbb{R}^{n\times n}$ such that $\|W\| \leq \bar{W}$, solutions of (11.14) initialized in $B_n(0,\rho)$ under the output feedback $u^* = \eta(S(x); l, \delta, \gamma)$ defined in (11.15) verify $x \in [\underline{R}, \bar{R}]$ and $|x - \gamma| \leq \epsilon$ when $t \geq T$.*

Using the output feedback introduced above, for any solution x initialized in $B_n(0,\rho)$, x is ensured to enter the set Ω_δ in finite time by taking $l > 0$ sufficiently large. Then, by analysis of the evolution of $e^2 := |x - \gamma|^2$ we are ensured that x gets arbitrarily close to γ when l is sufficiently large.

We recall the following

Definition 11.7. *A vector valued function $g \in C(\mathbb{R}, \mathbb{R}^n)$ verifies the Persistent Excitation condition (PE) if there exist $a, b, \tau > 0$ such that*

$$aI_n \preccurlyeq \int_{t-\tau}^{t} g(s)g(s)^\top ds \preccurlyeq bI_n. \tag{11.17}$$

If there exists $U \subset \mathbb{R}^n$ such that $g \in C(\mathbb{R}, U)$, then we write $g \in \mathrm{PE}(U)$.

The following corollary immediately follows from the previous lemma

Corollary 11.8. *Let $T, \rho, \bar{W}, L > 0$ and $0 < \delta < 1/6$. Let $\gamma \in C^{1,L}\left(\mathbb{R}, \Omega_\delta^n\right)$ such that $S(\gamma) \in \mathrm{PE}\left((0,1)^n\right)$. There exists an application $\Phi_t^* : C([0,t], \mathbb{R}^n) \to \mathbb{R}^{n\times n} \times \mathbb{R}^n$ and $l^* > 0$ such that for any $l > l^*$, for any $W \in \mathbb{R}^{n\times n}$ such that $\|W\| \le \bar{W}$, for any solution of (11.14) with initial in $B_n(0, \rho)$ and output feedback u^* in to the form of (11.15), we have*

$$\Phi_t^*\left(y|_{[0,t]}\right) = (W, x(t))\,, \quad \forall\, t \ge T. \tag{11.18}$$

Proof. Let $0 < T_1 < T_2 < T$. First, from the second point of Lemma 11.6, under the output feedback u^*, we know that $x(t) \in (\underline{R}, \bar{R})$ for all times $t \ge T_1$ for a high enough $l > 0$. We have therefore $x = S^{-1}(y)$. Using now Proposition 11.14 we can retrieve $WS(x(t))$ for all time $t \ge T_2$ using the sliding mode observer for $z_1 = x$, $z_2 = WS(x)$. To ease the proof we consider it as a virtual output $y_1 = WS(x(t))$. From Lemma 11.6, choosing γ such that $S(\gamma) \in \mathrm{PE}((0,1)^n)$, it is immediate from the regularity on S that for a high enough $l > 0$, $S(x(\cdot)) \in \mathrm{PE}((0,1)^n)$ after $t \ge T$. Therefore

$$W = \left(\int_{t-\tau}^{t} y_1(s)y(s)^\top\right)\left(\int_{t-\tau}^{t} y(s)y(s)^\top ds\right)^{-1}, \tag{11.19}$$

for all times $t \ge T$, with $\tau > 0$ small enough which can be tuned from the choice of γ. □

11.5 IDENTIFICATION OF A 2-LAYER SYSTEM

In this section, we consider the identification problem for the following 2-level system

$$\begin{cases} \dot{x}_1 = -x_1 + \psi_1(x_1, t) + WS(x_2), \\ \dot{x}_2 = -x_2 + \psi_2(x_1, x_2, t) + u_2, \\ y = (y_1, y_2) = (x_1, \psi_1(x_1(t), t))\,, \end{cases} \tag{11.20}$$

where $(x_1, x_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, $W \in \mathbb{R}^{d_1\times d_2}$ is left invertible, $\psi_1 \in C_b(\mathbb{R}^{d_1} \times [0, +\infty), \mathbb{R}^{d_1})$, $\psi_2 \in C_b(\mathbb{R}^{d_2} \times \mathbb{R}^{d_1} \times [0, +\infty), \mathbb{R}^{d_2})$ and $u_2 \in C_b([0, +\infty), \mathbb{R}^{d_2})$. Solutions of (11.20) are global in time and for any $R \ge R_0 := \|\psi_1\|_\infty + \|\psi_2\|_\infty + \|u\|_\infty > 0$ the ball $B_{d_1+d_2}(0, R)$ is an invariant set for (11.20).

11.5.1 *Identification of interaction terms*

We now want to retrieve from the output y in equation (11.20) the matrix W. To this aim, given two times $t_2 > t_1 > 0$, we consider the map $\Psi_{t_1}^{t_2} : C_b([0,T], \mathbb{R}^{2d_1}) \to \mathbb{R}^{d_1}$ that associates to an output $y = (y_1, y_2)$, defined up to a time $T > t_2$, the vector

$$\Psi_{t_1}^{t_2}(y) := y_1(t_2) - e^{-t_2+t_1} y_1(t_1) - e^{-t_2} \int_{t_1}^{t_2} e^s y_2(s) ds. \tag{11.21}$$

From Lemma 11.4 we have the following.

Lemma 11.9. *Let $t_2 > t_1 > 0$, $\rho > 0$, $m > 0$, and $w \in \mathcal{A}_n := \{0,1\}^n \subset \mathbb{R}^n$. Then, there exists a constant control $u_2 \in C_b([0,t_2], \mathbb{R}^{d_2})$, depending only on (t_1, t_2, ρ, m, w), such that, for any ψ_1 and ψ_2 such that $\|\psi_2\|_\infty \leq m$, for any solutions of (11.20) initialized in $B_{d_1}(0,\rho) \times B_{d_2}(0,\rho)$ we have that*

$$\Psi_{t_1}^{t_2}(y) = (1 - e^{-(t_2-t_1)}) W w. \tag{11.22}$$

Proof. Using Lemma 11.4 for the chosen w, t_1, ρ, m, there exists $\mu > 0$, such that choosing u_2 to be a constant control as defined in (7.22), we have that for any solutions $x = (x_1, x_2)$ of (11.20) initialized in $B_{d_1}(0,\rho) \times B_{d_2}(0,\rho)$, $S(x_2(t)) = w$ for all times $t \geq t_1$. Using then Duhammel's formula for the equation on x_1, we obtain indeed that (11.21) is equal to (11.22). □

Applying the previous lemma multiples times in a row, we have the following corollary.

Corollary 11.10. *Let $T > 0$ and consider two sequences t_1^k and t_2^k such that*

$$0 < \ldots < t_1^k < t_2^k < t_1^{k+1} < \ldots < t_2^{d_2} = T, \tag{11.23}$$

and $t_2^k - t_1^k = \alpha$ for some constant $\alpha > 0$, and any $k \in [\![1, d_2]\!]$. Let $\rho > 0$, and $m > 0$. Then, there exists a control $u_2 \in C_b([0,T], \mathbb{R}^{d_2})$, depending only on $(T, (t_1^k)_k, (t_2^k)_k, \rho, m)$, such that, for any solution of (11.20) initialized in $B_{d_1}(0,\rho) \times B_{d_2}(0,\rho)$, letting $\Psi_k := \Psi_{t_1^k}^{t_2^k}$, we have

$$W = \frac{1}{(1-e^{-\alpha})} \begin{bmatrix} \Psi_1(y) & \Psi_2(y) & \cdots & \Psi_{d_2}(y) \end{bmatrix}.$$

11.5.2 *Estimation of $S(x_2)$*

We consider system (11.20). At first, we want to estimate $S(x_2)$ from the knowledge of the output y and the matrix W. In this case ψ_2 remains unknown.

To achieve this, it suffices to apply a finite time observer as the one introduced in Appendix 11.9. This yields the following

Proposition 11.11. *Let $T > 0$, $\rho > 0$. For any $m, \bar{u} > 0$ such that*

$$\|\psi_2\|_\infty \leq m, \quad \|u\|_\infty \leq \bar{u}, \tag{11.24}$$

there exists Ψ_1, depending only on $(T, \rho, m, \bar{u})$, such that for any solutions of (11.20) initialized in $B_{d_1}(0,\rho) \times B_{d_2}(0,\rho)$ with output y, we have

$$\Psi_1(W, y|_{[0,t]}) = S(x_2(t)), \quad \forall\, t \geq T. \tag{11.25}$$

Combining all the previous results of this section, the following stands proved.

Theorem 11.12. *Assume ψ_2 is Lipschitz continuous. Let $T, \rho > 0$ and $m > 0$ such that $\|\psi_2\|_\infty \leq m$. Then, there exists a family of maps $U_t : C_b([0,t], \mathbb{R}^{d_1}) \to \mathbb{R}^{d_2}$ and map $\Phi_t : C_b([0,t], \mathbb{R}^{d_1}) \to \mathbb{R}^{d_1 \times d_2} \times \mathbb{R}^{d_2} \times \mathbb{R}^{d_2})$ such that for any solution of (11.20) with initial conditions in $B_{d_1+d_2}(0,\rho)$ and output feedback control $u_2(t) = U_t(y|_{[0,t]})$, we have*

$$\Phi_t \left(y|_{[0,t]} \right) = (W, S(x_2(t))) , \quad \forall\, t \geq T. \tag{11.26}$$

11.6 FINITE TIME STATE ESTIMATION WITH KNOWN PARAMETERS

In this section, we establish that the state of any k-layer subsystem $V^k = (V_1, \cdots, V_k)$ of system (11.1) can be estimated in finite from V_1 when the parameters Θ_k are known. More precisely, we demonstrate that this can be achieved using output feedback controls of the form (11.15). This result is particularly relevant for subsequent developments, as the same class of controls will be employed for the identification process.

Recall the definition of Θ_k in (11.4) and Ω_δ in (11.16). We have the following.

Lemma 11.13. *Let $k \in [\![2, p]\!]$. Let $T > 0$ and consider any sequence of times $0 = T_0 < T_1 < T_2 < \cdots < T_k = T$. Let $\rho > 0$, $0 < \delta < 1/6$. For any sequence $\gamma_i \in C^{1,L}(\mathbb{R}, \Omega_\delta^{n_i})$ with $i \in [\![2, k]\!]$, for any $\Theta_k \in \mathbb{R}^{Q_k}$, there exists a family of maps $O_t^i : \mathbb{R}^{Q_i} \times C([0,t], \mathbb{R}^{n_1}) \to \mathbb{R}^{n_i}$ and a map $O_t^* : \mathbb{R}^{Q_k} \times C([0,t], \mathbb{R}^{n_1}) \to \mathbb{R}^{N_k}$, and $l^* > 0$, such that for any $l > l^*$, for any $0 \leq \alpha \leq 1$, letting for $i \in [\![2, k]\!]$*

$$u_i = \eta(\alpha \mathbf{1}_{t \leq T_i} + O_t^i \mathbf{1}_{t \geq T_i}; l, \delta, \gamma_i), \tag{11.27}$$

we have for any solution of (11.1) with starting in $B_N(0,\rho)$,

$$O_t^i(\Theta_i, y|_{[0,t]}) = S(V_i(t)), \quad \forall t \geq T_{i-1}, \tag{11.28}$$

$$O_t^*(\Theta_k, y|_{[0,t]}) = (V_1(t), \cdots, V_k(t)), \quad t \geq T. \tag{11.29}$$

Proof. The proof relies on a cascade of sliding mode observers (11.32) as introduced in Appendix 11.9, and follows a recursive argument. We show here how to estimate V_2 from V_1, and $S(V_3)$ from V_2. The rest of the proof follows.

From the output V_1, using $z_1 = V_1$ and $z_2 = W_{12}S(V_2)$, we satisfy the conditions of Proposition 11.14 for a time $T_1 > 0$. For all $t \geq 0$, we obtain an estimate $\hat{z}_2(t)$ such that $\hat{z}_2(t) = W_{12}S(V_2(t))$ for all $t \geq T_1$. Since W_{12} is known and left-invertible by Assumption 11.2, applying its left inverse $W_{11}^\dagger$ to the estimate allows us to construct the first mapping

$$O_t^2(\Theta_1, y|_{[0,t]}) = S(V_2(t)), \quad t \geq T_1. \tag{11.30}$$

From Lemma 11.6, we know that under such feedbacks, there exists a sufficiently large $l_1 > 0$ such that for a duration of $(T_2 - T_1)/2$, the trajectory V_2 remains in the invertible region of S for all $t \geq T_1 + (T_2 - T_1)/2$. Setting $y_1(t) = S^{-1}(O_t^1(\Theta_1, y|_{[T_0,t]}))$, it follows that $y_1(t) = V_2(t)$

for all $t \geq T_1 + (T_2 - T_1)/2$. Following the same process, we can tune a sliding mode observer such that $\bar{O}_t^3$ and O_t^3 have

$$O_t^3(\Theta_2, y|_{[0,T]}) = \bar{O}_t^3(\Theta_2, y_1|_{[T_2,t]}) = S(V_3(t)), \quad t \geq T_2.$$

Using this argument recursively yields the desired result. The map O_t^* is obtained by combining all the O_t^i and S^{-1}. □

11.7 PROOF OF THEOREM 11.3

Let $0 < T_1 < T_2 < T_3 < T$. Let $L > 0$. From the structure of (11.1), the parameters Θ_{k+1} appear in the dynamic of $V^{k+1} = (V_1, \cdots, V_{k+1})$. Corollary 11.5 ensures that there exist constant controls u_{q+i},...,u_p, such that after time T_1, all lower layers $q \geq k+2$ will be functionally disconnected. The state V^{k+1} will follow a $k+1$-layers system.

Lemma 11.13 ensure that for $\rho > 0$ high enough, $0 < \delta < 1/6$, $\gamma_i \in C^{1,L}(\mathbb{R}, \Omega^{n_i})$ and for the time $T_2 - T_1 > 0$, placing feedback controls u_i of type (11.27), there exist $l_1 > 0$ such that for any $l \geq l_1$ we can recover the values of the vector $V^k = (V_1, \cdots, V_k)$ for all times after $t \geq T_2$. We denote for simplicity of proof, $y_1(t) := V^k(t)$.

We now wish to retrieve $W_{k,k+1}$. We consider the dynamics of V^k and V_{k+1} as a two layer system (11.20) with

$$\psi_1(t) = \sum_{j=1}^{n_k} W_{k,j} S(V_j(t)), \; \psi_2(t) = \sum_{j=1}^{n_{k+1}} W_{k+1,j} S(V_j(t)).$$

From assumption, the matrices $W_{k,j}$ are known and from the virtual output y_1 we indeed have the same form of output as in (11.20). Moreover, we also verify the condition of ψ_2 unknown with known bound $(k+1)\bar{W}$.

We now are in condition to use Theorem 11.12 for a time $T_3 - T_2 > 0$, $\rho_2 > 0$ large enough and $m = k\bar{W}$. The theorem ensures that there exist an adequate piecewise continuous control $u_{k+1} : [T_2, T_3] \to \mathbb{R}^{n_k}$ such that we can recover the values of $W_{k,k+1}$ and $S(V_{k+1}(t))$ for all times $t \geq T_3$. We consider, once again for simplicity, $y_2(t) = S(V_{k+1}(t))$ as a virtual output our system after time T_2. Recall, from Lemma 11.13, that estimation of (y_1, y_2) is maintained thanks to the output feedback controls u_i of class (11.27), for $i \in [\![1, k]\!]$, and the fact that the values of $S(V_{k+1}(t))$ can be recovered through a sliding mode observer independently of the control u_{k+1} once $W_{k,k+1}$ is known.

Let now $\gamma = (\gamma_1, \gamma_2, \cdots, \gamma_{k+1})$, with $\gamma_i \in C^{1,L}(\mathbb{R}, \Omega^{n_i})$ be such that $S(\gamma) \in \mathrm{PE}\left((0,1)^n\right)$. We are in condition to apply Corollary 11.8 for a time $T - T_3 > 0$, ρ_3 high enough and the chosen γ. The theorem ensures that we recover the block matrix W after time $T > 0$. In particular, we can recover its last row $W_{k+1,j}$ for $j \in [\![1, k+1]\!]$. Moreover, the corollary also yields that we can recover $V^k(t)$ after time T. Theorem 11.3 therefore stands proved. □

11.8 ILLUSTRATION

In this section, we provide an illustration of the method's principle by highlighting how saturations are exploited to reveal the weights. We consider a toy example of a two-layer system with $n_1 = 2$, $n_2 = 1$. We pick arbitrary matrices

$$W_{11} = \begin{pmatrix} 0.6 & -0.2 \\ 0.4 & -0.7 \end{pmatrix}, \qquad W_{12} = \begin{pmatrix} 0.3 \\ -0.9 \end{pmatrix},$$

$$W_{21} = \begin{pmatrix} -0.4 & -0.1 \end{pmatrix}, \qquad W_{22} = \begin{pmatrix} -0.9 \end{pmatrix}.$$

Regarding the saturation function S, we pick $\underline{R} = 0$, $\bar{R} = 1$ and $S(x) = x$ on $[0, 1]$. Then we define the dates $T_1 = 1$, $T_2 = 3$, $T_3 = 5$, $T_4 = 8$. We set the controls to be

$$(u_1^1, u_1^2, u_2) = \begin{cases} (0, 0, -3) & \text{on } [0, T_1), \\ (3, -3, -3) & \text{on } [T_1, T_2), \\ (-3, 3, -3) & \text{on } [T_2, T_3), \\ (-3, -3, 3) & \text{on } [T_3, T_4). \end{cases}$$

Simulating this system yields Figure 11.1.

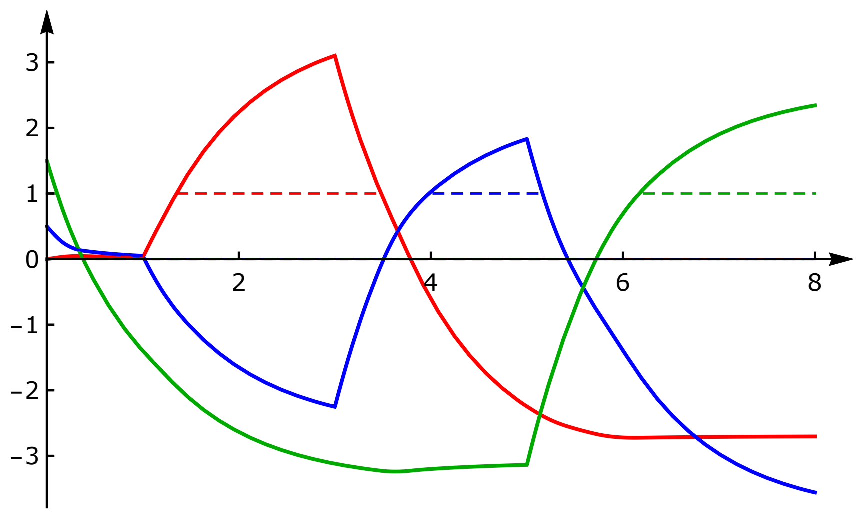


Figure 11.1: Trajectory of the toy system. In red and blue are represented (x_1^1, x_1^2) of the first layer, while in green is represented x_2 of the second layer. Dashed are the trajectories saturated at 1.

On $[0, T_3)$, $u_2 = -3$ in order for $S(x_2)$ to be 0 on $[T_1, T_3)$. On $[T_1, T_2)$ and $[T_2, T_3)$, controls are picked so that $S(x_1) = (1, 0)$ and $S(x_1) = (0, 1)$ on $[T_2 - 1, T_2)$ and $[T_3 - 1, T_3)$ respectively. Likewise, on $[T_3, T_4)$, controls are picked so that $S(x_1) = (0, 0)$ and $S(x_2) = 1$ on $[T_4 - 1, T_4)$. As a result, the virtual measurement $z = \dot{x}_1 + x_1 - u_1$, that we can estimate (following from Section 11.5 and Appendix 11.9), occupies on $[T_k - 1, T_k)$, $k = 2, 3, 4$, $(W_{11}^{11}, W_{11}^{21})$, $(W_{11}^{12}, W_{11}^{22})$ and W_{12}, respectively. See Figure 11.2.

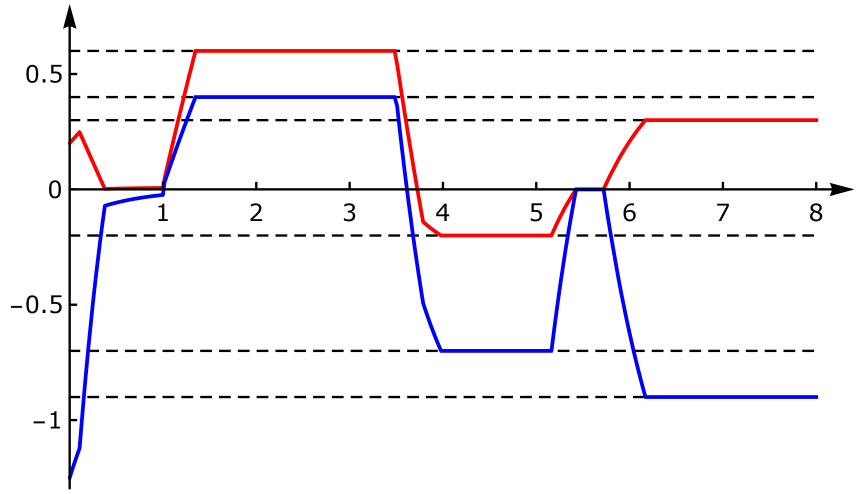


Figure 11.2: Values of the virtual measurement $z = \dot{x}_1 + x_1 - u_1$. In red and blue are represented (z^1, z^2), respectively. Dashed lines are the constant weights in W_{11}, W_{12}. For $k = 2, 3, 4$, z and coincides with the weights over $[T_k - 1, T_k)$, which allows through estimation of z to recover them.

11.9 CONCLUSION

In this paper, we established novel identification methods for systems governed by networked Wilson-Cowan-like equation. By leveraging the properties of saturations, we designed open-loop controls that allow extraction of the weight parameters in finite time. In single-layer systems, self-referential weights can be identified by steering the system into specific configurations. For two-layer systems, where the top layer is measured, similar principles enable the recovery of layer-interaction weights. Furthermore, we demonstrated how well-designed feedback mechanisms can facilitate a descent into the lower layer, effectively allowing us to treat it as a virtual measurement. Extending this approach, we recursively generalized our results to systems with an arbitrary number of layers.

Additional parameters are identifiable within this framework, but fall beyond the scope of this paper. These ideas warrant further exploration, particularly with a view to relaxing certain assumptions on the structure of the system.

In the realm of artificial neural networks, this type of scheme would be well fitted to identify weights resulting from learning, as proposed in [10]. By performing identification on a neural network with substantially fewer nodes than the one providing the measure, this scheme could be used to produce a distillation of the inital one (see, e.g., [119]).

APPENDIX: FINITE-TIME OBSERVERS

Let us recall some well known results on homogeneous/sliding mode [116, 7]. Consider the following system with $z = (z_1, z_2) \in \mathbb{R}^n \times \mathbb{R}^n$,

$$\begin{cases} \dot{z}_1 = g_1(z_1, t) + z_2, \\ \dot{z}_2 = g_2(z_1, z_2, t), \\ y = z_1. \end{cases} \tag{11.31}$$

with g_i continuous maps. Sliding mode observers allow to recover the state of the system in finite time, without explicit knowledge of g_2 other than a bound. More precisely, consider the following observer, depending on $L, K_1, K_2 > 0$:

$$\begin{cases} \dot{\hat{z}}_1 = g_1(y, t) + \hat{z}_2 + LK_1 \lceil y_1 - \hat{z}_1 \rfloor^{\frac{1}{2}} \\ \dot{\hat{z}}_2 \in L^2 K_2 \mathrm{sign}(y_1 - \hat{z}_1). \end{cases} \tag{11.32}$$

Here, for $\sigma \in \mathbb{R}$, we let $\lceil \sigma \rfloor^{\frac{1}{2}} := \mathrm{sign}(\sigma)\sqrt{|\sigma|}$, with $\mathrm{sign}(\sigma)$ the set valued mapping

$$\mathrm{sign}(\sigma) := \begin{cases} \{1\} & \text{if } \sigma > 0, \\ [-1, 1] & \text{if } \sigma = 0, \\ \{-1\} & \text{if } \sigma < 0. \end{cases}$$

We give the following proposition, adapted to our context from [26, Proposition 4].

Proposition 11.14. *Assume that g_1 and g_2 piecewise continuously differentiable and bounded, and that $z_1 \mapsto g_1(z_1, t)$ is $\frac{1}{2}$-Hölder continuous uniformly w.r.t. t. Let $m > 0$ be such that $\|g_2\|_\infty < m$. Then, there exist $K_1, K_2 \in \mathbb{R}$ such that for any time $T > 0$ and $\rho > 0$ there exist a gain $L^* > 0$ such that for any $L > L^*$ and for any $\hat{z}(0), z(0) \in B_{2n}(0, \rho)$ we have,*

$$\hat{z}(t) = z(t), \quad \forall\, t \geq T. \tag{11.33}$$

BIBLIOGRAPHY

[1] L. F. Abbott and C. van Vreeswijk. Asynchronous states in networks of pulse-coupled oscillators. *Physical Review E*, 48(2):1483–1490, 1993.

[2] L. Albantakis, C. Bernard, N. Brenner, E. Marder, and R. Narayanan. The brain's best kept secret is its degenerate structure. *Journal of Neuroscience*, 44(40), 2024.

[3] S.-i. Amari. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological cybernetics*, 27(2):77–87, 1977.

[4] S.-i. Amari. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological cybernetics*, 27(2):77–87, 1977.

[5] V. Andrieu, D. Astolfi, and P. Bernard. Observer design via interconnections of second-order mixed sliding-model/linear differentiators. *International Journal of Robust and Nonlinear Control*, 31(9):3631–3657, 2021.

[6] V. Andrieu and L. Praly. On the existence of a kazantzis-kravaris/luenberger observer. *SIAM Journal on Control and Optimization*, 45(2):432–456, 2006.

[7] V. Andrieu, L. Praly, and A. Astolfi. Homogeneous approximation, recursive observer design, and output feedback. *SIAM Journal on Control and Optimization*, 47(4):1814–1850, 2008.

[8] A. M. Annabi, J.-B. Pomet, D. Prandi, and L. Sacchelli. Activity estimation via distributed measurements in an orientation sensitive neural fields model of the visual cortex. *Mathematics of Control, Signals, and Systems*, 37(3):737–767, 2025.

[9] A. M. Annabi, J.-B. Pomet, D. Prandi, and L. Sacchelli. Identification of saturated networked systems. In *2025 IEEE 64th Conference on Decision and Control (CDC)*, pages 4208–4213. IEEE, 2025.

[10] T. Aoyama and H. Ichikawa. Reconstruction of weight matrices in neural networks: A method of correlating outputs with inputs. *Chemical and pharmaceutical bulletin*, 39(5):1222–1228, 1991.

[11] S. Aranovskiy, A. Bobtsov, R. Ortega, and A. Pyrkin. Performance enhancement of parameter estimators via dynamic regressor extension and mixing procedure. *IEEE Transactions on Automatic Control*, 62(7):3546–3550, 2017.

[12] M. Arcak and P. Kokotović. Nonlinear observers: a circle criterion design and robustness analysis. *Automatica*, 37(12):1923–1930, 2001.

[13] L. Astolfi, F. Cincotti, D. Mattia, F. D. V. Fallani, S. Salinari, M. Marciani, C. Wilke, A. Doud, H. Yuan, B. He, et al. Estimation of the cortical activity from simultaneous multi-subject recordings during the prisoner's dilemma. In *2009 annual international conference of the ieee engineering in medicine and biology society*, pages 1937–1939. IEEE, 2009.

[14] D. Avitabile, G. J. Lord, and K. Meddouni. State and parameter estimation for a neural model of local field potentials. *arXiv preprint arXiv:2512.07842*, 2025.

[15] F. A. Azevedo, L. R. Carvalho, L. T. Grinberg, J. M. Farfel, R. E. Ferretti, R. E. Leite, W. J. Filho, R. Lent, and S. Herculano-Houzel. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *Journal of comparative neurology*, 513(5):532–541, 2009.

[16] F. Babiloni, F. Cincotti, M. Marciani, S. Salinari, L. Astolfi, A. Tocci, F. Aloise, F. Fallani, S. Bufalari, D. Mattia, et al. The estimation of cortical activity for brain-computer interface: applications in a domotic context. *Computational Intelligence and Neuroscience*, 2007, 2007.

[17] A. S. Bazanella, M. Gevers, and L. s. Miˇ sković. Closed-loop identification of MIMO systems: a new look at identifiability and experiment design. *Eur. J. Control*, 16(3):228–239, 2010.

[18] A. S. Bazanella, M. Gevers, and L. Miškovic. Closed-loop identification of mimo systems: a new look at identifiability and experiment design. *European Journal of Control*, 16(3):228–239, 2010.

[19] R. Ben-Yishai, R. L. Bar-Or, and H. Sompolinsky. Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences*, 92(9):3844–3848, 1995.

[20] C. Bernard. Brain's best kept secret: Degeneracy. *Eneuro*, 10(11), 2023.

[21] P. Bernard. *Observer design for nonlinear systems*, volume 479. Springer, 2019.

[22] P. Bernard and V. Andrieu. Luenberger observers for nonautonomous nonlinear systems. *IEEE Transactions on Automatic Control*, 64(1):270–281, 2018.

[23] P. Bernard, V. Andrieu, and D. Astolfi. Observer design for continuous-time dynamical systems. *Annual Reviews in Control*, 2022.

[24] P. Bernard, V. Andrieu, and L. Praly. Expressing an observer in preferred coordinates by transforming an injective immersion into a surjective diffeomorphism. *SIAM Journal on Control and Optimization*, 56(3):2327–2352, 2018.

[25] P. Bernard and L. Marconi. Hybrid implementation of observers in plant's coordinates with a finite number of approximate inversions and global convergence. *Automatica*, 111:108654, 2020.

[26] P. Bernard, L. Praly, and V. Andrieu. Observers for a non-lipschitz triangular form. *Automatica*, 82:301–313, 2017.

[27] M. Bertalmio, L. Calatroni, V. Franceschi, B. Franceschiello, A. Gomez Villa, and D. Prandi. Visual illusions via neural dynamics: Wilson–cowan-type models and the efficient representation principle. *Journal of neurophysiology*, 123(5):1606–1618, 2020.

[28] M. Bertalmio, L. Calatroni, V. Franceschi, B. Franceschiello, and D. Prandi. Cortical-inspired wilson–cowan-type equations for orientation-dependent contrast perception modelling. *Journal of Mathematical Imaging and Vision*, 63:263–281, 2021.

[29] G. Besançon. *Nonlinear observers and applications*, volume 363. Springer, 2007.

[30] G. Besanöon, Q. Zhang, and H. Hammouri. High-gain observer based state and parameter estimation in nonlinear systems. *IFAC Proceedings Volumes*, 37(13):327–332, 2004.

[31] G. Besançon. Nonlinear observers and applications. In *Lecture Notes in Control and Information Sciences*, volume 363, pages 211–222. Springer, 2007.

[32] A. Björner. *Oriented matroids*. Number 46. Cambridge University Press, 1999.

[33] B. Blumenfeld, D. Bibitchkov, and M. Tsodyks. Neural network model of the primary visual cortex: from functional architecture to lateral connectivity and back. *Journal of computational neuroscience*, 20:219–241, 2006.

[34] W. H. Bosking, Y. Zhang, B. Schofield, and D. Fitzpatrick. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *Journal of neuroscience*, 17(6):2112–2127, 1997.

[35] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*, volume 15 of *SIAM Studies in Applied Mathematics*. SIAM, 1994.

[36] M. Breakspear. Dynamic models of large-scale brain activity. *Nature neuroscience*, 20(3):340–352, 2017.

[37] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, New York, 3rd edition, 2008.

[38] P. C. Bressloff. Spatiotemporal dynamics of continuum neural fields. *Journal of Physics A: Mathematical and Theoretical*, 45(3):033001, 2011.

[39] L. Brivadis, V. Andrieu, P. Bernard, and U. Serres. Further remarks on kkl observers. *Systems & Control Letters*, 172:105429, 2023.

[40] L. Brivadis, A. Chaillet, and J. Auriol. Online estimation of hilbert-schmidt operators and application to kernel reconstruction of neural fields. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pages 597–602. IEEE, 2022.

[41] L. Brivadis, A. Chaillet, and J. Auriol. Adaptive observer and control of spatiotemporal delayed neural fields. *Systems & Control Letters*, 186:105777, 2024.

[42] N. Brunel. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of computational neuroscience*, 8(3):183–208, 2000.

[43] T. B. Burghi, T. O'Leary, and R. Sepulchre. Distributed online estimation of biophysical neural networks. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pages 628–634. IEEE, 2022.

[44] T. B. Burghi and R. Sepulchre. Adaptive observers for biophysical neuronal circuits. *IEEE Transactions on Automatic Control*, 69(8):5020–5033, 2023.

[45] E. Busvelle and J. Gauthier. High-gain and non-high-gain observers for nonlinear systems. In *Contemporary trends in nonlinear geometric control theory and its applications*, pages 257–286. World Scientific, 2002.

[46] C. Cakan, N. Jajcay, and K. Obermayer. neurolib: A simulation framework for whole-brain neural mass modeling. *Cognitive Computation*, 15(4):1132–1152, 2023.

[47] R. Carron, A. Chaillet, A. Filipchuk, W. Pasillas-Lépine, and C. Hammond. Closing the loop of deep brain stimulation. *Frontiers in systems neuroscience*, 7:112, 2013.

[48] M. Chong, R. Postoyan, D. Nešić, L. Kuhlmann, and A. Varsavsky. A robust circle criterion observer with application to neural mass models. *Automatica*, 48(11):2986–2989, 2012.

[49] G. Chowdhary and E. Johnson. Concurrent learning for convergence in adaptive control without persistency of excitation. In *49th IEEE Conference on Decision and Control (CDC)*, pages 3674–3679. IEEE, 2010.

[50] G. Chowdhary, M. Muehlegg, and E. Johnson. Exponential parameter and tracking error convergence guarantees for adaptive controllers without persistency of excitation. *International Journal of Control*, 87(8):1583–1603, 2014.

[51] B. J. Cook, A. D. Peterson, W. Woldman, and J. R. Terry. Neural field models: A mathematical overview and unifying framework. *Mathematical Neuroscience and Applications*, 2, 2022.

[52] S. Coombes, P. beim Graben, R. Potthast, and J. Wright. *Neural fields: theory and applications*. Springer, 2014.

[53] S. Coombes, P. beim Graben, R. Potthast, and J. Wright. A tutorial on neural field theory. *Neural fields*, pages 1–43, 2014.

[54] S. Coombes, Y. M. Lai, M. Şayli, and R. Thul. Networks of piecewise linear neural mass models. *European Journal of Applied Mathematics*, 29(5):869–890, 2018.

[55] J.-M. Coron. *Control and nonlinearity*. Number 136. American Mathematical Soc., 2007.

[56] O. David and K. J. Friston. A neural mass model for meg/eeg:: coupling and neuronal dynamics. *NeuroImage*, 20(3):1743–1755, 2003.

[57] E. B. Davies. *Heat Kernels and Spectral Theory*. Cambridge University Press, Cambridge, 1989.

[58] G. Deco, V. K. Jirsa, and A. R. McIntosh. Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nature Reviews Neuroscience*, 12(1):43–56, 2011.

[59] G. Deco, V. K. Jirsa, and A. R. McIntosh. Resting brains never rest: computational insights into potential cognitive architectures. *Trends in Neurosciences*, 36(5):268–274, 2013.

[60] G. Deco, V. K. Jirsa, P. A. Robinson, M. Breakspear, and K. Friston. The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLoS computational biology*, 4(8):e1000092, 2008.

[61] A. Destexhe and T. J. Sejnowski. The Wilson-Cowan model, 36 years later. *Biological Cybernetics*, 101(1):1–2, 2009.

[62] G. I. Detorakis, A. Chaillet, S. Palfi, and S. Senova. Closed-loop stimulation of a delayed neural fields model of parkinsonian stn-gpe network: a theoretical and computational study. *Frontiers in neuroscience*, 9:237, 2015.

[63] M. Dewar, K. Scerri, and V. Kadirkamanathan. Data-driven spatio-temporal modeling using the integro-difference equation. *IEEE Transactions on Signal Processing*, 57(1):83–91, 2009.

[64] F. Deza, E. Busvelle, J. Gauthier, and D. Rakotopara. High gain estimation for nonlinear systems. *Systems & control letters*, 18(4):295–299, 1992.

[65] D. M. Dunstan, P. Abbott, J. R. Terry, and R. Garnett. Global nonlinear approach for mapping parameters of neural mass models. *PLoS Computational Biology*, 19(3):e1010985, 2023.

[66] D. Efimov and T. Raïssi. Design of interval observers for uncertain dynamical systems. *Automation and Remote Control*, 77(2):191–225, 2016.

[67] D. Efimov, T. Raïssi, S. Chebotarev, and A. Zolghadri. Interval state observer for nonlinear time varying systems. *Automatica*, 49(1):200–205, 2013.

[68] G. T. Einevoll, C. Kayser, N. K. Logothetis, and S. Panzeri. Modelling and analysis of local field potentials for studying the function of cortical circuits. *Nature Reviews Neuroscience*, 14(11):770–785, 2013.

[69] H. Elliott and W. Wolovich. Parameter adaptive identification and control. *IEEE Transactions on Automatic Control*, 24(4):592–599, 1979.

[70] S. Emelyanov, S. Korovin, S. Nikitin, and M. Nikitina. Observers and output differentiators for nonlinear systems. *Doklady Akademii Nauk*, 306:556–560, 1989.

[71] K.-J. Engel and R. Nagel. *One-Parameter Semigroups for Linear Evolution Equations*. Springer, New York, 2000.

[72] L. Escuain-Poole, E. Cano, P. S. Garcia, E. Ros, and E. Pereda. Extracranial estimation of neural mass model parameters using the unscented kalman filter. *Frontiers in Applied Mathematics and Statistics*, 4:46, 2018.

[73] F. Esfandiari and H. Khalil. Observer-based design of uncertain systems: recovering state feedback robustness under matching conditions. In *Proceedings of Allerton Conference, Monticello, IL*, pages 97–106, 1987.

[74] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, Providence, RI, 2nd edition, 2010.

[75] X. Fan and M. Arcak. Observer design for systems with multivariable monotone nonlinearities. *Systems & Control Letters*, 50(4):319–330, 2003.

[76] L. Farina and S. Rinaldi. *Positive linear systems: theory and applications*. John Wiley & Sons, 2011.

[77] M. Farza, M. M'Saad, T. Maatoug, and M. Kamoun. Adaptive observers for nonlinearly parameterized class of nonlinear systems. *Automatica*, 45(10):2292–2299, 2009.

[78] A. F. Filippov. *Differential equations with discontinuous righthand sides: control systems*, volume 18. Springer Science & Business Media, 2013.

[79] J. Fix, M. Geist, O. Pietquin, and H. Frezza-Buet. Dynamic neural field optimization using the unscented kalman filter. In *2011 IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB)*, pages 1–7. IEEE, 2011.

[80] T. Floquet and J.-P. Barbot. Super twisting algorithm-based step-by-step sliding mode observers for nonlinear systems with unknown inputs. *International journal of systems science*, 38(10):803–815, 2007.

[81] G. B. Folland. *Real analysis: modern techniques and their applications*. John Wiley & Sons, 1999.

[82] W. J. Freeman. Mass action in the nervous system. *New York: Academic*, 1975.

[83] D. R. Freestone, P. Aram, M. Dewar, K. Scerri, D. B. Grayden, and V. Kadirkamanathan. A data-driven framework for neural field modeling. *NeuroImage*, 56(3):1043–1058, 2011.

[84] A. Galka, T. Ozaki, H. Muhle, U. Stephani, and M. Siniatchkin. A data-driven model of the generation of human EEG based on a spatially distributed stochastic wave equation. *Cognitive Neurodynamics*, 2(2):101–113, 2008.

[85] J. Gauthier and G. Bornard. Observability for any u(t) of a class of non-linear systems. *IEEE Transactions on Automatic Control*, 26(4):922–926, 1981.

[86] J.-P. Gauthier, H. Hammouri, and S. Othman. A simple observer for nonlinear systems applications to bioreactors. *IEEE Transactions on automatic control*, 37(6):875–880, 1992.

[87] J.-P. Gauthier and I. Kupka. *Deterministic observation theory and applications*. Cambridge University Press, Cambridge, 2001.

[88] J.-P. Gauthier and I. Kupka. *Deterministic observation theory and applications*. Cambridge University Press, 2001.

[89] A. Ghosh, Y. Rho, A. R. McIntosh, R. Kötter, and V. K. Jirsa. Noise during rest enables the exploration of the brain's dynamic repertoire. *PLoS computational biology*, 4(8):e1000196, 2008.

[90] M. Giaccagli, V. Andrieu, S. Tarbouriech, and D. Astolfi. Lmi conditions for contraction, integral action, and output feedback stabilization for a class of nonlinear systems. *Automatica*, 154:111106, 2023.

[91] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer, Berlin, 2nd edition, 2001.

[92] M. Goodfellow, K. Schindler, and G. Baier. Self-organised transients in a neural mass model of epileptogenic tissue dynamics. *NeuroImage*, 59(3):2644–2660, 2012.

[93] G. Goodwin and K. Sin. *Adaptive Filtering, Prediction and Control*. Prentice Hall, 1984.

[94] J.-L. Gouzé, A. Rapaport, and M. Z. Hadj-Sadok. Interval observers for uncertain biological systems. *Ecological Modelling*, 133(1-2):45–56, 2000.

[95] M. Grüter and K.-O. Widman. The Green function of a uniformly elliptic equation in arbitrarily shaped domains. *Manuscripta Mathematica*, 37(3):303–342, 1982.

[96] R. Gulbinaite, B. İlhan, and R. VanRullen. The triple-flash illusion reveals a driving role of alpha-band reverberations in visual perception. *Journal of Neuroscience*, 37(30):7219–7230, 2017.

[97] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, V. J. Wedeen, and O. Sporns. Mapping the structural core of human cerebral cortex. *PLOS Biology*, 6(7):e159, 2008.

[98] A. Hartoyo, P. J. Cadusch, D. T. Liley, and D. G. Hicks. Parameter estimation and identifiability in a neural population model for electro-cortical activity. *PLoS computational biology*, 15(5):e1006694, 2019.

[99] J. A. Herron, M. C. Thompson, T. Brown, H. J. Chizeck, J. G. Ojemann, and A. L. Ko. Cortical brain–computer interface for closed-loop deep brain stimulation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(11):2180–2187, 2017.

[100] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500–544, 1952.

[101] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500, 1952.

[102] C. J. Honey, O. Sporns, L. Cammoun, X. Gigandet, J.-P. Thiran, R. Meuli, and P. Hagmann. Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences*, 106(6):2035–2040, 2009.

[103] X.-S. Hu, K.-S. Hong, S. S. Ge, and M.-Y. Jeong. Kalman estimator-and general linear model-based on-line brain activation mapping by near-infrared spectroscopy. *Biomedical engineering online*, 9(1):1–15, 2010.

[104] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160(1):106, 1962.

[105] J. D. Jackson. *Classical Electrodynamics*. John Wiley & Sons, New York, 3rd edition, 1999.

[106] B. H. Jansen and V. G. Rit. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological Cybernetics*, 73(4):357–366, 1995.

[107] V. K. Jirsa and H. Haken. Field theory of electromagnetic brain activity. *Physical review letters*, 77(5):960, 1996.

[108] V. K. Jirsa and H. Haken. The cortical wave equation: modeling and applications. *Handbook of Brain Connectivity*, pages 31–64, 2010.

[109] E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, and A. J. Hudspeth. *Principles of Neural Science*. McGraw-Hill, 5th edition, 2013.

[110] H. Khalil and L. Praly. High-gain observers in nonlinear feedback control. *International Journal of Robust and Nonlinear Control*, 24, 2013.

[111] H. K. Khalil and L. Praly. High-gain observers in nonlinear feedback control. *International Journal of Robust and Nonlinear Control*, 24(6):993–1015, 2014.

[112] A. J. Krener and A. Isidori. Linearization by output injection and nonlinear observers. *Systems & Control Letters*, 3(1):47–52, 1983.

[113] M. V. Kulikova and G. Y. Kulikov. Data-driven parameter estimation in stochastic dynamic neural field models. *Digital Signal Processing*, 139:104206, 2023.

[114] J. Kybic, M. Clerc, T. Abboud, and O. Faugeras. Neural mass model parameter identification for meg/eeg. In *Proceedings of SPIE*, 2007.

[115] G. A. Leonov, D. V. Ponomarenko, and V. B. Smirnova. *Frequency-domain methods for nonlinear analysis: theory and applications*, volume 9. World Scientific, 1996.

[116] A. Levant. Robust exact differentiation via sliding mode technique. *automatica*, 34(3):379–384, 1998.

[117] A. Levant. Higher-order sliding modes and arbitrary-order exact robust differentiation. In *Proceedings of the European Control Conference*, pages 996–1001, 2001.

[118] A. Levant. Higher-order sliding modes, differentiation and output-feedback control. *International Journal of Control*, 76(9-10):924–941, 2003.

[119] Y. Li, L. Liu, G. Wang, Y. Du, and P. Chen. EGNN: Constructing explainable graph neural networks via knowledge distillation. *Knowledge-Based Systems*, 2022.

[120] X. Liu, D. Miao, and Q. Gao. A robust nonlinear observer for a class of neural mass models. *The Scientific World Journal*, 2014(1):215943, 2014.

[121] L. Ljung. *System Identification: Theory for the User*. Prentice Hall, 2nd edition, 1999.

[122] A. López-Cuevas, B. Castillo-Toledo, L. Medina-Ceja, and C. Ventura-Mejía. State and parameter estimation of a neural mass model from electrophysiological signals during the status epilepticus. *NeuroImage*, 113:374–386, 2015.

[123] R. Marino, L. Buffoni, L. Chicchi, F. D. Patti, D. Febbe, L. Giambagli, and D. Fanelli. Learning in wilson-cowan model for metapopulation. *Neural Computation*, 37(4):701–741, 2025.

[124] R. Marino and P. Tomei. Global adaptive observers and output-feedback stabilization for a class of nonlinear systems. In *Foundations of Adaptive Control*, volume 160 of *Lecture Notes in Control and Information Sciences*, pages 455–493. Springer, Berlin, Heidelberg, 1991.

[125] V. Markounikau, C. Igel, A. Grinvald, and D. Jancke. A dynamic neural field model of mesoscopic cortical activity captured with voltage-sensitive dye imaging. *PLoS computational biology*, 6(9):e1000919, 2010.

[126] F. Mazenc and O. Bernard. Interval observers for linear time-invariant systems with disturbances. *Automatica*, 47(1):140–147, 2011.

[127] L. Menini, C. Possieri, and A. Tornambe. A "practical" observer for nonlinear systems. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 3015–3020. IEEE, 2017.

[128] R. Moran, D. A. Pinotsis, and K. J. Friston. Neural masses and fields in dynamic causal modeling. *Frontiers in Computational Neuroscience*, 7:57, 2013.

[129] J. A. Moreno and M. Osorio. A lyapunov approach to second-order sliding mode controllers and observers. In *2008 47th IEEE conference on decision and control*, pages 2856–2861. IEEE, 2008.

[130] O. Nelles. Nonlinear system identification. *Measurement Science and Technology*, 13(4):646–646, 2002.

[131] C. Novara, T. Vincent, K. Hsu, M. Milanese, and K. Poolla. Parametric identification of structured nonlinear systems. *Automatica J. IFAC*, 47(4):711–721, 2011.

[132] R. Ortega, S. Aranovskiy, A. A. Pyrkin, A. Astolfi, and A. A. Bobtsov. New results on parameter estimation via dynamic regressor extension and mixing: Continuous and discrete-time cases. *IEEE Transactions on Automatic Control*, 66(5):2265–2272, 2020.

[133] K. O'Hashi, T. Fekete, T. Deneux, R. Hildesheim, C. Van Leeuwen, and A. Grinvald. Interhemispheric synchrony of spontaneous cortical states at the cortical column level. *Cerebral Cortex*, 28(5):1794–1807, 2018.

[134] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer, New York, 1983.

[135] D. A. Pinotsis and K. J. Friston. On conductance-based neural field models. *Frontiers in Computational Neuroscience*, 7:158, 2013.

[136] R. Postoyan and A. Chaillet. L'automatique au service des neurosciences. *La Recherche*, 2020.

[137] R. Postoyan, M. S. Chong, D. Nešić, and L. Kuhlmann. Parameter and state estimation for a class of neural mass models. In *51st IEEE Conference on Decision and Control (CDC)*, pages 2236–2241. IEEE, 2012.

[138] C. Qian. A homogeneous domination approach for global output feedback stabilization of a class of nonlinear systems. In *Proceedings of the 2005, American Control Conference, 2005.*, pages 4708–4715 vol. 7, 2005.

[139] T. Raïssi, D. Efimov, and A. Zolghadri. Interval state estimation for a class of nonlinear systems. *IEEE Transactions on Automatic Control*, 57(1):260–265, 2012.

[140] R. Rajamani. Observers for lipschitz nonlinear systems. *IEEE Transactions on Automatic Control*, 43(3):397–401, 1998.

[141] L. Ramos, E. Di Meglio, V. Morgenthaler, L. da Silva, and P. Bernard. Numerical design of luenberger observers for nonlinear systems. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 5435–5442. IEEE, 2020.

[142] P. Sanz-Leon, S. A. Knock, A. Spiegler, and V. K. Jirsa. The virtual brain: a simulator of primate brain network dynamics. *Frontiers in neuroinformatics*, 7:10, 2013.

[143] S. J. Schiff. *Neural control engineering: the emerging intersection between control theory and neuroscience*. MIT press, 2011.

[144] S. J. Schiff and T. Sauer. Kalman filter control of a model of spatiotemporal cortical dynamics. *Journal of Neural Engineering*, 5(1):1–8, 2008.

[145] S. H. Scott. Optimal feedback control and the neural basis of volitional motor control. *Nature Reviews Neuroscience*, 5(7):532–546, 2004.

[146] H. L. Smith. *Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems*. American Mathematical Society, 1995.

[147] T. Söderström and P. Stoica. *System Identification*. Prentice Hall International Series in Systems and Control Engineering. Prentice Hall, 1989.

[148] E. D. Sontag. *Mathematical control theory: deterministic finite dimensional systems*, volume 6. Springer Science & Business Media, 2013.

[149] M. Sood, P. Besson, M. Muthalib, U. Jindal, S. Perrey, A. Dutta, and M. Hayashibe. Nirs-eeg joint imaging during transcranial direct current stimulation: online parameter estimation with an autoregressive model. *Journal of neuroscience methods*, 274:71–80, 2016.

[150] E. Sorrell, M. E. Rule, and T. O'Leary. Brain–machine interfaces: Closed-loop control in an adaptive system. *Annual Review of Control, Robotics, and Autonomous Systems*, 4:167–189, 2021.

[151] N. P. Subramaniyam, C. Cakan, I. Bojak, M. Breakspear, and V. K. Jirsa. Sensitivity-analysis-guided bayesian parameter estimation for neural mass models. *Physical Review E*, 110(4):044208, 2024.

[152] C. Tamekue, D. Prandi, and Y. Chitour. Reproducibility via neural fields of visual illusions induced by localized stimuli. *arXiv preprint arXiv:2401.09108*, 2024.

[153] J. R. Terry, W. Woldman, A. D. Peterson, and B. J. Cook. Neural field models: A mathematical overview and unifying framework. *Mathematical Neuroscience and Applications*, 2, 2022.

[154] F. E. Thau. Observing the state of non-linear dynamic systems. *International Journal of Control*, 17(3):471–479, 1973.

[155] S. Tilwani and C. O'Reilly. Benchmarking deep jansen-rit parameter inference, 2024.

[156] E. Todorov and M. I. Jordan. Optimal feedback control as a theory of motor coordination. *Nature neuroscience*, 5(11):1226–1235, 2002.

[157] A. Tornambè. Use of asymptotic observers having-high-gains in the state and parameter estimation. In *Proceedings of the 28th IEEE Conference on Decision and Control,*, pages 1791–1794. IEEE, 1989.

[158] E. Trélat. *Contrôle optimal: théorie & applications*, volume 36. Vuibert Paris, 2005.

[159] A. van der Schaft and H. Schumacher. *An introduction to hybrid dynamical systems*, volume 251 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag London, Ltd., London, 2000.

[160] R. Veltz. *Nonlinear analysis methods in neural field models*. PhD thesis, Paris Est, 2011.

[161] R. Veltz and O. Faugeras. Illusions in the ring model of visual orientation selectivity. *arXiv preprint arXiv:1007.2493*, 2010.

[162] R. Veltz and O. Faugeras. Local/global analysis of the stationary solutions of some neural field equations. *SIAM Journal on Applied Dynamical Systems*, 9(3):954–998, 2010.

[163] T. L. Vincent, C. Novara, K. Hsu, and K. Poolla. Input design for structured nonlinear system identification. *Automatica*, 46(6):990–998, 2010.

[164] T. L. Vincent, C. Novara, K. Hsu, and K. Poolla. Input design for structured nonlinear system identification. *Automatica J. IFAC*, 46(6):990–998, 2010.

[165] F. Wendling, F. Bartolomei, J. J. Bellanger, and P. Chauvel. Epileptic fast activity can be explained by a model of impaired gabaergic dendritic inhibition. *European Journal of Neuroscience*, 15(9):1499–1508, 2002.

[166] F. Wendling, J.-J. Bellanger, F. Bartolomei, and P. Chauvel. Relevance of nonlinear lumped-parameter models in the analysis of depth-eeg epileptic signals. *Biological cybernetics*, 83(4):367–378, 2000.

[167] H. Wilson and J. Cowan. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, 13(2):55–80, 1973.

[168] H. R. Wilson and J. D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12(1):1–24, 1972.

[169] T. Zaslavsky. *Facing up to arrangements: Face-count formulas for partitions of space by hyperplanes: Face-count formulas for partitions of space by hyperplanes*, volume 154. American Mathematical Soc., 1975.

[170] A. Zemouche and M. Boutayeb. On lmi conditions to design observers for lipschitz nonlinear systems. *Automatica*, 49(2):585–591, 2013.

[171] Q. Zhang. Revisiting different adaptive observers through a unified formulation. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 3067–3072. IEEE, 2005.

[172] K. Åström and B. Wittenmark. *Adaptive Control*. Dover Publications, 2nd edition, 2008.

A

TOOLS AND BACKGROUND ON TRIANGULAR OBSERVER DESIGN

A.1 INTRODUCTION

The design of asymptotic observers for nonlinear dynamical systems remains a challenging problem in control theory. While numerous approaches have been developed over the past decades, most successful designs are restricted to specific classes of systems that can be transformed into particular normal forms. Among these, triangular forms occupy a central position due to their structural properties that enable systematic observer design.

As noted in the comprehensive review by Bernard et al. [21], the observer design problem typically involves finding a transformation that converts the original system dynamics into a form where observer design becomes tractable. For a significant class of nonlinear systems, this transformation leads to triangular structures where the system dynamics exhibit a cascade form that proves particularly amenable to observer design.

This chapter focuses on triangular systems and their associated observer methodologies. We begin by introducing the general triangular form and discussing the conditions under which a nonlinear system can be transformed into this structure. We then explore the main observer design techniques developed for triangular systems, particularly high-gain observers and homogeneous observers, along with the specific conditions required on the system nonlinearities to ensure convergence.

A.2 CANONICAL FORMS AND ASSOCIATED OBSERVABILITY NOTIONS

In nonlinear systems theory, the observer design problem has been successfully solved for a limited number of canonical forms. The general methodology involves finding a transformation that maps the original system dynamics into one of these canonical forms, where established observer designs can be applied [21].

Consider the general nonlinear system:

$$\dot{x} = f(x,t), \quad y = h(x,t) \tag{A.1}$$

where $x \in \mathbb{R}^{n_x}$ is the state, $y \in \mathbb{R}^{n_y}$ is the output, and $f : \mathbb{R}^{n_x} \times \mathbb{R} \to \mathbb{R}^{n_x}$, $h : \mathbb{R}^{n_x} \times \mathbb{R} \to \mathbb{R}^{n_y}$ are sufficiently smooth functions.

This transformation approach is formalized through a mapping $T : \mathbb{R}^{n_x} \times \mathbb{R} \to \mathbb{R}^{n_z}$ that satisfies the fundamental partial differential equation:

$$\frac{\partial T}{\partial x}(x,t)f(x,t) + \frac{\partial T}{\partial t}(x,t) = F(T(x,t), h(x,t), t) \tag{A.2}$$

where the specific structure of $F : \mathbb{R}^{n_z} \times \mathbb{R}^{n_y} \times \mathbb{R} \to \mathbb{R}^{n_z}$ defines the target canonical form. Each canonical form is associated with a particular observability notion that guarantees the existence and injectivity of the transformation.

A.2.1 *Triangular Form and Differential Observability*

The triangular canonical form represents one of the most widely studied and practically successful structures for observer design [88]. This form is characterized by its cascade structure that naturally accommodates observer design through high-gain or homogeneous approaches.

For autonomous systems, the transformation to triangular form is typically constructed using the output and its successive Lie derivatives:

$$T : \mathbb{R}^{n_x} \to \mathbb{R}^{n_z}, \quad T(x) = \begin{pmatrix} h(x) \\ L_f h(x) \\ \vdots \\ L_f^{n_z-1} h(x) \end{pmatrix} \tag{A.3}$$

with $n_z \geq n_x$, where $L_f h(x)$ denotes the Lie derivative of h along f.

The existence and injectivity of this transformation are guaranteed under the condition of strong differential observability:

Definition A.1 (Strong Differential Observability [85])**.** *The system* (A.1) *is strongly differentially observable of order n_z on a set $\mathcal{X} \subset \mathbb{R}^{n_x}$ if the mapping $T(x)$ defined in* (A.3) *is injective on $\mathcal{X}$ and has full rank n_x for all $x \in \mathcal{X}$.*

This notion requires that the state can be uniquely and instantaneously determined from the output and its first $n_z - 1$ time derivatives. Under this condition, the system can be transformed into the triangular form [88]:

$$\begin{aligned} \dot{z}_1 &= z_2 + \phi_1(z_1, t) \\ \dot{z}_2 &= z_3 + \phi_2(z_1, z_2, t) \\ &\vdots \\ \dot{z}_{n_z} &= \phi_{n_z}(z_1, \ldots, z_{n_z}, t) \\ y &= z_1 \end{aligned} \tag{A.4}$$

The transformation satisfies equation (A.2) with $F(z, y, t)$ having the triangular structure where each component F_i depends only on $z_1, \ldots, z_{i+1}$. For time-varying systems, the transformation becomes $T(x, t)$ and the required observability notion extends to uniform differential observability, where the injectivity holds uniformly in time.

A.2.2 *State-Affine Form and Linearizability by Output Injection*

The state-affine canonical form represents another important class where observer design benefits from linear systems theory [86]. In this form, the system appears linear in the state but nonlinear in the output, enabling the use of Kalman-like observers.

For the state-affine canonical form, the transformation $T : \mathbb{R}^{n_x} \times \mathbb{R} \to \mathbb{R}^{n_z}$ must satisfy:

$$\frac{\partial T}{\partial x}(x,t)f(x,t) + \frac{\partial T}{\partial t}(x,t) = A(h(x,t),t)T(x,t) + \varphi(h(x,t),t) \tag{A.5}$$

where $A : \mathbb{R}^{n_y} \times \mathbb{R} \to \mathbb{R}^{n_z \times n_z}$ and $\varphi : \mathbb{R}^{n_y} \times \mathbb{R} \to \mathbb{R}^{n_z}$.

The existence of such a transformation is guaranteed under the condition of linearizability by output injection:

Definition A.2 (Linearizability by Output Injection [112])**.** *The system (A.1) is locally linearizable by output injection around a point x_0 if there exists a neighborhood U of x_0 and a diffeomorphism $T : U \to \mathbb{R}^{n_z}$ with $n_z = n_x$ such that in the coordinates $z = T(x)$, the system takes the form:*

$$\dot{z} = Az + \varphi(y,t), \quad y = Cz \tag{A.6}$$

with (A, C) an observable pair.

This property requires that the observation space $\mathcal{O}$, defined as the smallest linear space containing $h(x)$ and closed under Lie differentiation along the vector fields of the system, has finite dimension. This represents a very strong observability condition, as it essentially requires that the nonlinear system behaves like a linear system from the observation perspective [112].

A.2.3 *Hurwitz Form and Backward Distinguishability*

The Hurwitz form, central to Kazantzis-Kravaris/Luenberger (KKL) observers, offers a very general framework for nonlinear observer design [6]. This form enables observer design through a simple copy of the dynamics in transformed coordinates.

For the Hurwitz form, the transformation $T : \mathbb{R}^{n_x} \times \mathbb{R} \to \mathbb{R}^{n_z}$ satisfies:

$$\frac{\partial T}{\partial x}(x,t)f(x,t) + \frac{\partial T}{\partial t}(x,t) = \Lambda T(x,t) + \Gamma h(x,t) \tag{A.7}$$

where $\Lambda \in \mathbb{R}^{n_z \times n_z}$ is a Hurwitz matrix and $\Gamma \in \mathbb{R}^{n_z \times n_y}$.

The existence and injectivity of this transformation are guaranteed under the condition of backward distinguishability. A system is backward distinguishable if for any two distinct initial states, there exists some time in the past where the corresponding outputs differ. This represents a weaker observability notion compared to differential observability, as it only requires distinguishability over time intervals rather than instantaneously.

Definition A.3 (Backward Distinguishability [22])**.** *The system (A.1) is backward distinguishable on a set $\mathcal{X} \subset \mathbb{R}^{n_x}$ in time $t_d > 0$ if for any $t_0 \geq t_d$ and any pair of solutions $x_a(\cdot), x_b(\cdot)$ with $x_a(t_0) \neq x_b(t_0)$ both in $\mathcal{X}$, there exists $s \in [t_0 - t_d, t_0]$ such that:*

$$h(x_a(s), s) \neq h(x_b(s), s) \tag{A.8}$$

This represents a weaker observability notion compared to differential observability, as it only requires that states can be distinguished over time intervals rather than instantaneously. Under backward distinguishability, the resulting canonical form is [6, 22]:

$$\dot{z} = \Lambda z + \Gamma y, \quad y = H(z,t) \tag{A.9}$$

where $H(T(x,t),t) = h(x,t)$.

A.2.4 *Comparative Analysis and Practical Considerations*

The relationship between these observability notions forms a strict hierarchy [21]:

$$\begin{aligned}\text{Differential observability} &\Rightarrow \text{Linearizability by output injection}\\ &\Rightarrow \text{Backward distinguishability}\end{aligned}$$

Each implication is strict, with differential observability representing the strongest condition and backward distinguishability the weakest. This hierarchy reflects the trade-off in nonlinear observer design: stronger observability conditions enable simpler canonical forms and more powerful observer designs, but are satisfied by fewer systems.

The triangular form, enabled by differential observability, has proven particularly valuable in practice [88]. While it requires the strongest observability condition, this condition is often verifiable for physical systems, and the resulting observers offer strong convergence properties with relatively straightforward implementation. The cascade structure of triangular forms naturally accommodates both high-gain designs for Lipschitz nonlinearities and homogeneous designs for more general nonlinearities.

In contrast, the KKL approach based on backward distinguishability applies to the broadest class of systems but often requires solving challenging PDEs for the transformation and may lead to higher-dimensional observers [6]. The state-affine form strikes a middle ground but imposes structural constraints that limit its applicability to specific classes of nonlinearities [112].

The choice of canonical form thus depends on both the system's inherent observability properties and the practical constraints of implementation, with the triangular form often providing the most favorable balance for engineering applications [88, 21].

A.3 OBSERVER DESIGN FOR TRIANGULAR SYSTEMS

Each canonical form comes with its own practical constraints for observer implementation. Linearization by output injection, while offering a clear observer structure, applies only to a strict class of systems with specific structural properties, limiting its broader applicability. Similarly, the KKL approach provides an elegant observer design where the entire complexity is shifted to computing the transformation T and its pseudo-inverse - challenges that often prove formidable in practice due to the need to solve partial differential equations and implement real-time inversion algorithms.

In contrast, differential observability naturally leads to systems of the form:

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = x_3, \quad \dots, \quad \dot{x}_n = \phi(x_1, \dots, x_n) \tag{A.10}$$

where the observer design benefits from the chain of integrators structure. However, many physical systems admit more general transformations through domain-specific knowledge, yielding generalized triangular forms:

$$\dot{x}_i = x_{i+1} + \phi_i(x_1, \dots, x_i), \quad i = 1, \dots, n-1, \quad \dot{x}_n = \phi_n(x_1, \dots, x_n) \tag{A.11}$$

where nonlinearities appear throughout the cascade rather than being confined to the final equation.

This generalized triangular structure accommodates a wider range of physical systems while maintaining the essential cascade property that enables systematic observer design. Crucially, unlike the clear-cut designs for KKL and state-affine forms, the observer methodology for triangular systems is not universal but depends fundamentally on the regularity properties of the nonlinear functions ϕ_i. The specific observer design - whether high-gain, homogeneous, or interconnected - must be tailored to the Lipschitz continuity, Hölder continuity, or other regularity characteristics of these nonlinearities.

It is precisely this flexibility to accommodate various nonlinearity types through customized observer designs, combined with the practical availability of transformations to triangular form for many physical systems, that makes this approach the focus of our subsequent discussion on practical observer implementation.

A.3.1 *High-Gain Observers*

High-gain observers represent one of the most successful approaches for triangular systems. Their development can be traced to several early works in the late 1980s [70, 157, 73, 64]. The fundamental idea involves using high gains to dominate the nonlinear terms and achieve fast convergence.

For a system in the triangular form (A.11), a typical high-gain observer takes the form [110]:

$$\begin{aligned}
\dot{\hat{z}}_1 &= \hat{z}_2 + \phi_1(\hat{z}_1, t) + \ell k_1 (y - H(\hat{z}_1, t)) \\
\dot{\hat{z}}_2 &= \hat{z}_3 + \phi_2(\hat{z}_1, \hat{z}_2, t) + \ell^2 k_2 (y - H(\hat{z}_1, t)) \\
&\vdots \\
\dot{\hat{z}}_{n_z} &= \phi_{n_z}(\hat{z}_1, \dots, \hat{z}_{n_z}, t) + \ell^{n_z} k_{n_z} (y - H(\hat{z}_1, t))
\end{aligned} \tag{A.12}$$

where $\ell > 0$ is the high-gain parameter and $k_1, \dots, k_{n_z}$ are chosen such that the matrix:

$$A - KC = \begin{pmatrix} -k_1 & 1 & & 0 \\ \vdots & & \ddots & \\ & & & 1 \\ -k_{n_z} & & \dots & 0 \end{pmatrix} \tag{A.13}$$

is Hurwitz, with A and C having the canonical form of (A.4).

The key feature of this design is that for sufficiently large ℓ, the observer achieves uniform exponential stability in the z-coordinates. The convergence rate can be made arbitrarily fast by increasing ℓ, though this comes at the cost of increased sensitivity to measurement noise, a phenomenon known as the "peaking phenomenon" [110].

The sufficient conditions for convergence typically require that the functions ϕ_i satisfy global Lipschitz conditions with respect to their arguments. More precisely, assuming the triangular form (A.11) with $H(z_1, t) = z_1$ for simplicity, we require that there exist positive constants $\underline{a}, \overline{a}, \overline{b}$ such that for all (z, y, t):

$$\underline{a} < \left| \frac{\partial F_i}{\partial z_{i+1}}(z, y, t) \right| < \overline{a}, \quad \left| \frac{\partial F_i}{\partial z_j}(z, y, t) \right| < \overline{b} \quad \text{for } j \leq i \tag{A.14}$$

where $F_i(z, y, t) = z_{i+1} + \phi_i(z_1, \ldots, z_i, t)$ for $i = 1, \ldots, n_z - 1$, and $F_{n_z}(z, y, t) = \phi_{n_z}(z_1, \ldots, z_{n_z}, t)$.

Under these conditions, there exists $\ell^* \geq 1$ such that for all $\ell > \ell^*$, the observer (A.12) is a tunable asymptotic observer. The estimation error in the z-coordinates satisfies [88]:

$$|z(t) - \hat{z}(t)| \leq c\ell^{n_z - 1} \exp(-\ell\lambda t)|z(0) - \hat{z}(0)|, \quad \forall t \tag{A.15}$$

for some constants c, λ independent of ℓ.

A.3.2 *Homogeneous Observers*

When the triangular system does not satisfy global Lipschitz conditions, homogeneous observers provide an alternative approach. These observers, sometimes referred to as sliding-mode observers in specific cases, can handle nonlinearities with lower regularity requirements.

The development of homogeneous observers can be traced to the work of Levant [117] and has been extensively studied [7, 26]. For a system in triangular form with constant A and C matrices in the canonical form of (A.4), a homogeneous observer takes the form:

$$\dot{\hat{z}} = A\hat{z} + \phi(\hat{z}, t) + \mathcal{D}_\ell k_{\text{Hom}}(y - C\hat{z}) \tag{A.16}$$

where $\mathcal{D}_\ell = \text{diag}(\ell, \ell^2, \ldots, \ell^{n_z})$ is the high-gain scaling matrix, and k_{Hom} is a homogeneous correction term.

The homogeneous correction terms of degree $\mathrm{r} \in [-1, \frac{1}{n_z - 1})$ are defined component-wise as [5]:

$$k_{\text{Hom},j}(s) = \mu_j \,\text{sign}(s)|s|^{\frac{r_j}{r_1}}, \quad r_j = 1 - \mathrm{r}(n_z - j) \tag{A.17}$$

where μ_j are gains to be tuned, and $\text{sign}(s)$ is the set-valued sign function.

For $\mathrm{r} = -1$ and $n_z = 2$, we recover the well-known super-twisting differentiator:

$$k_{\text{Hom},1}(s) = \mu_1 \,\text{sign}(s)|s|^{\frac{1}{2}}, \quad k_{\text{Hom},2}(s) = \mu_2 \,\text{sign}(s) \tag{A.18}$$

The key advantage of homogeneous observers is their ability to handle nonlinearities that satisfy less restrictive conditions than global Lipschitz requirements. Specifically, if the functions ϕ_i satisfy for all $i \in \{1, \ldots, n_z\}$ and for all (y, z_a, z_b, t) [26]:

$$|\phi_i(y, z_{2a}, \ldots, z_{ia}, t) - \phi_i(y, z_{2b}, \ldots, z_{ib}, t)| \leq \overline{b} \sum_{j=2}^{i} |z_{ja} - z_{jb}|^{\alpha_{ij}} \tag{A.19}$$

with

$$\frac{1 + (n_z - i - 1)\mathrm{r}}{1 + (n_z - j)\mathrm{r}} \leq \alpha_{ij} < 1, \quad j \leq i \leq n_z \tag{A.20}$$

then there exists $\mathrm{r} < 0$ and $\ell^* > 0$ such that for all $\ell > \ell^*$, the homogeneous observer achieves finite-time convergence.

In the limiting case where the lower bounds in (A.20) are equalities, we take $\mathrm{r} = -1$ and no restriction is imposed on ϕ_{n_z} besides boundedness, recovering the robustness properties of sliding-mode observers.

A.3.3 *Interconnected Observers*

To address limitations of both high-gain and homogeneous observers, interconnected structures have been proposed. These cascades of lower-dimensional observers can improve performance in terms of peaking phenomenon and sensitivity to measurement noise while relaxing the conditions on the nonlinearities.

A typical interconnected observer takes the form [5]:

$$\begin{aligned} \dot{\hat{z}}_{ii} &= F_i(y, t, \hat{z}_{12}, \hat{z}_{23}, \ldots, \hat{z}_{i(i+1)}) - \ell_i k_{i,1}(\hat{z}_{ii} - \hat{z}_{(i-1)1}) \\ \dot{\hat{z}}_{i(i+1)} &= F_{i+1}(y, t, \hat{z}_{12}, \hat{z}_{23}, \ldots, \hat{z}_{(i+1)(i+2)}) - \ell_i^2 k_{i,2}(\hat{z}_{ii} - \hat{z}_{(i-1)1}) \\ i &= 0, \ldots, n_z - 1 \end{aligned} \tag{A.21}$$

with the convention $\hat{z}_{01} = y$ and $\hat{z}_{n_z(n_z+1)} = 0$.

This structure increases the relative degree between the measurement noise and the state estimates, improving filtering properties at high frequencies. When the gains $k_{i,j}$ are selected following a mixed linear/homogeneous approach, the interconnected observer can handle nonlinearities satisfying affine incremental bounds [5]:

$$|\phi_i(y, z_{2a}, \ldots, z_{ia}, t) - \phi_i(y, z_{2b}, \ldots, z_{ib}, t)| \leq \bar{b}_0 + \bar{b}_1 \sum_{j=2}^{i} |z_{ja} - z_{jb}| \tag{A.22}$$

which are less restrictive than the homogeneous conditions (A.19).

A.4 ON THE INVERSE OF THE TRANSFORMATION T

Designing an observer in transformed coordinates solves only part of the state estimation problem. The crucial step of recovering the state estimate in the original coordinates requires computing the left-inverse $\mathcal{T}$ of the transformation T, which presents significant theoretical and practical challenges that often determine the feasibility of the overall observer implementation [21].

In most practical implementations, the left-inverse $\mathcal{T}$ is computed through real-time optimization, solving the minimization problem:

$$\hat{x}(t) = \arg\min_{x \in \mathcal{X}} |T(x, t) - \hat{z}(t)|^2 \tag{A.23}$$

This approach appears straightforward but faces fundamental limitations that stem from the non-convex nature of the cost function. The objective function $J(x) = |T(x,t) - \hat{z}(t)|^2$ typically exhibits multiple local minima when T is nonlinear, leading to convergence to incorrect state estimates [127].

The fundamental theoretical condition ensuring the existence of a well-behaved left-inverse $\mathcal{T}$ is the uniform injectivity of T. This property guarantees that the transformation preserves distinctness of states in a quantitatively controlled manner. Formally, T is uniformly injective if there exists a $\mathcal{K}$-function w such that:

$$|x_a - x_b| \leq w(|T(x_a,t) - T(x_b,t)|) \quad \forall x_a, x_b \in \mathcal{X}, t \geq \bar{t} \tag{A.24}$$

This condition ensures the existence of a left-inverse $\mathcal{T}$ satisfying the crucial continuity property:

$$|\mathcal{T}(\hat{z}) - \mathcal{T}(z)| \leq w(|\hat{z} - z|) \quad \forall \hat{z}, z \in T(\mathcal{X}) \tag{A.25}$$

For time-dependent transformations, this extends to a $\mathcal{KL}$-function, ensuring that the convergence rate in the z-coordinates is preserved through the inversion process [21]. The uniform injectivity condition is both necessary and sufficient for guaranteeing that small estimation errors in the transformed coordinates translate to small errors in the original state estimates.

A central difficulty in implementing the left-inverse arises when T is an injective immersion but not surjective, which occurs frequently when $n_z > n_x$. In this case, the image $T(\mathbb{R}^{n_x})$ forms a proper submanifold of $\mathbb{R}^{n_z}$, and the observer state $\hat{z}$ may leave this submanifold during transients due to estimation errors or disturbances.

The fundamental challenge becomes defining $\mathcal{T}(\hat{z})$ for $\hat{z} \notin T(\mathbb{R}^{n_x})$. Indeed, observer trajectories frequently venture outside the image manifold, particularly during initial convergence or in the presence of measurement noise. The definition of $\mathcal{T}$ in these regions must be carefully constructed to ensure that the resulting state estimates remain meaningful [24].

A.4.1 *Advances in Inverse Computation*

Recent research has developed several strategies to address the inversion challenge, each with distinct advantages and limitations.

The diffeomorphic case, where T is a bijection with smooth inverse, offers the most straightforward implementation. Here, the observer can be designed directly in the original coordinates using the Jacobian:

$$\dot{\hat{x}} = \left(\frac{\partial T}{\partial x}(\hat{x})\right)^{-1} \mathcal{F}(T(\hat{x}), y, t) \tag{A.26}$$

This approach elegantly circumvents explicit inversion but requires careful monitoring to ensure the Jacobian remains invertible and that $\hat{x}$ stays within the domain of definition [24].

For the more common case where T is an immersion, Bernard et al. developed the method of dimension extension [24]. This approach constructs a surjective diffeomorphism T_e : $\mathbb{R}^{n_x} \times \mathbb{R}^{n_z - n_x} \to \mathbb{R}^{n_z}$ that extends T, allowing the observer to be implemented in the extended space. While theoretically powerful, this method requires careful construction of the extension and increases the system dimension.

A particularly promising direction is the hybrid approach developed by Bernard and Marconi [25], which employs a finite number of approximate inversions combined with reset strategies. This method guarantees global convergence while requiring only occasional computation of $\mathcal{T}$, making it suitable for applications where the inversion is computationally expensive. The key insight is to use an independent practical observer to reset the estimate when it approaches problematic regions, ensuring that the transformation inversion is only performed when reliable.

Other recent advances include learning-based approaches that approximate $\mathcal{T}$ offline using neural networks [141], though these methods introduce new challenges regarding certification of convergence and generalization guarantees.

A.4.2 *Conclusion*

The computation of the left-inverse $\mathcal{T}$ remains a critical challenge in nonlinear observer design, bridging the gap between theoretical existence results and practical implementation. The minimization approach, while conceptually simple, faces fundamental limitations due to non-convexity and local minima. The theoretical foundation provided by uniform injectivity ensures the existence of a well-behaved inverse, but the practical implementation requires careful consideration of how to define $\mathcal{T}$ outside the image manifold.

Recent advances in diffeomorphic implementations, dimension extension, and hybrid strategies with finite inversions offer promising directions for making nonlinear observer theory practically applicable across diverse engineering domains. The choice among these approaches depends on the specific properties of the transformation T and the computational constraints of the application, with each method offering distinct trade-offs between theoretical guarantees, computational complexity, and implementation simplicity [21, 25].

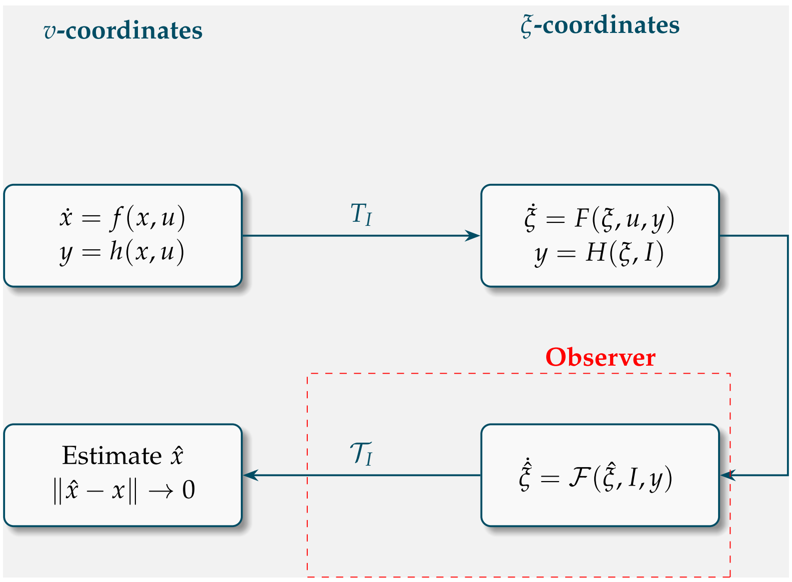


Figure A.1: Caption Describing the design of an observer