1. Introduction

1.1. Motivation

The knowledge of the internal state of a dynamical system is of crucial importance for many control applications, for example when stabilizing the system via state feedback, when monitoring compliance with safety-critical conditions or when detecting errors and external attacks. In most practical cases, however, the state cannot be completely measured for various (possibly physical or economic) reasons and therefore must be reconstructed using the available input-output signals. This is generally a challenging problem, especially in the presence of nonlinear systems and when robustness to model errors and measurement noise must be ensured.

Moving horizon estimation (MHE) [RMD20] is a modern optimization-based state estimation strategy that is naturally suitable for this purpose. Here, the current state estimate is obtained by solving an optimization problem involving a fixed number of past measurements, extracting the last state of the optimal estimated sequence, and repeating the online optimization in the next sampling time in a receding horizon fashion. It can be interpreted as an approximation to full information estimation (FIE), which optimizes over all available historical data. However, FIE is usually only of theoretical interest (particularly as a benchmark for MHE), since the complexity of the underlying optimization problem continuously grows with time and thus quickly becomes computationally intractable in practical applications.

MHE has several advantages over other state estimation methods: it is naturally applicable to nonlinear systems, provides the ability to include additional information such as constraints, is intuitive to tune, and yields optimal estimation results. Moreover, it is fairly easy to implement using high-level software packages (such as acados [Ver+21] and CasADi [And+18]), merely requiring knowledge of the model equations and corresponding computing resources. This is in strong contrast to most nonlinear observers; they require less computing power when applied, but the corresponding design is usually based on the search for a global transformation into a suitable observer normal form or the solution of a partial differential equation [BAA22], which is generally non-trivial and requires a deeper understanding of the underlying theory, representing a relatively large hurdle for use in practice. For these reasons, and not least because of the steadily growing availability of computing capacities and the development of highly efficient optimization algorithms, MHE is increasingly applied in various different fields, ranging from chemical and process engineering [HCE18; Els+21], mobile robotics and localization [Liu+17; Bre19], offshore engineering and freight transportation [CLH22], to medical applications [Kle+23].

However, the corresponding theory developed rather slowly, with little ability to provide practical tuning guidelines. Only recently, substantial progress has been made by deriving robustness properties of MHE under a relatively mild detectability condition, compare [RMD20]. Nevertheless, many problems and open questions remain that prevent the current MHE theory from providing any value beyond conceptual nature. Specifically, the following problems can be identified.

1. Practical relevance: The recently developed robustness guarantees for MHE require the knowledge of a particular detectability property for the design of the cost function, for which there is no systematic method for verification. Moreover, the results are mostly overly conservative, yielding unrealistic and practically irrelevant design guidelines and estimates on the horizon length.
2. Restrictive design: The most recent results in the field of nonlinear MHE focus on discrete-time systems and do not have a direct continuous-time counterpart. In this context, there is also a lack of fundamental theory on suitable continuous-time notions of detectability and robust stability. However, investigating corresponding MHE schemes is important, as the original physical system to be estimated usually corresponds to a continuous-time model. Having to discretize it first significantly complicates the system representation, restricts flexibility, can lead to additional discretization errors, and requires fixing a particular discretization scheme and sampling period beforehand.
3. Real-time capability: The computing power available in practice is often severely limited, and computing the global optimum at each time step is usually not possible within a fixed time interval. Instead, the solver is usually terminated with a suboptimal solution, which renders the theoretical guarantees invalid (as they usually depend on this criterion).
4. Parametric model uncertainties: In practical applications, the derived system model requires system identification and usually suffers from parametric uncertainties, as only noisy measurement data is available. This, however, may invalidate the available robustness guarantees, which crucially rely on an exact model of the system, or even cause the estimation error to become unstable. Adapting the parameters online to obtain a precise model is not directly possible, as it is yet unclear how to deal with potential lack of excitation (which often occurs frequently or unpredictably in practical applications).
5. Estimation performance: There is no general performance analysis of MHE available. It is therefore unclear how an MHE scheme must be designed in order to ultimately achieve a similar estimation performance to the (desired but impractical) FIE counterpart or a comparable benchmark.

This thesis aims at developing a deepened system-theoretic understanding of MHE and establishing desired robust stability and performance guarantees under realistic and practically relevant conditions, contributing to the greater goal of supplementing the great success of MHE in practical applications with a well-founded theory. In the following two sections, we provide an overview on the related literature and summarize the main contributions of this thesis.

1.2. Literature overview

In this section, we provide a brief overview over the literature related to the research topic. This of course covers works on nonlinear MHE, but also on system-theoretic properties such as detectability and robust stability, methods for combined state and parameter estimation in general, as well as performance and turnpike analysis in the context of optimal control.

1.2.1. Nonlinear detectability

While it is well understood how observability and detectability can be characterized and verified for linear systems (see, for example, Chapters 5 and 6 in [Son90]), this is not the case for general nonlinear systems. Here, one might transfer conceptually similar approaches and generally argue about the indistinguishability of different initial conditions based on the respective output signals or try to analyze the observation space of the system using Lie derivatives along the vector field (see, e.g., [Bes07, Ch. 1]). However, explicit verification of such rather abstract properties in practical applications is generally a complex and difficult problem.

A system-theoretic approach to characterize detectability for nonlinear systems is given by the concept of incremental input/output-to-state stability (i-IOSS). This property requires that the difference between any two trajectories of a dynamical system is upper bounded by the difference in their respective initial conditions, their inputs, and their outputs. Loosely speaking, if the differences between their respective inputs and outputs are small, then the difference between the states must also become small, which hence directly entails an indistinguishability property that is a natural characteristic of detectability in general.

The concept of i-IOSS was originally proposed in [SW97] to extend the notion of (non-incremental) input/output-to-state stability (IOSS)—which compares a system trajectory with the zero-trajectory and can thus only be regarded as "zero-detectability"—to a pair of arbitrary system trajectories. Introduced in an L^∞-to-L^∞ sense, it has been shown that a continuous-time system must necessarily satisfy the i-IOSS property to admit a robustly stable full-order state observer, and its discrete-time analogue has become the standard in the field of optimization-based state estimation, compare, for example, [RJ12; Ji+16; Mül17; RMD20; AR21; KM23; Sch+23; Hu24; Ale25].

The characterization of system properties via Lyapunov functions has turned out to be very useful for system analysis and the design of controllers and observers. Here, it is important to establish the equivalence between the Lyapunov function characterization and its corresponding original notion by means of converse theorems, in order to ensure that considering the Lyapunov function, which is usually easier when designing controllers and observers, is indeed without loss of generality. Such results are available (mostly in both continuous and discrete time) for, e.g., global asymptotic stability (GAS) in [LSW96] and [JW01], input-to-state stability (ISS) in [SW95] and [JW01], (non-incremental) IOSS in [KSW01] and [CT08], and integral

IOSS in [Ing01a; Ing01b]. Stronger, incremental notions are considered in [Ang02] and [TRK16], which address incremental GAS and incremental ISS (i-ISS), albeit under the condition that inputs and external signals (such as, e.g., time-varying parameters or disturbances) of the system take values in compact sets. The condition of compactness could be weakened by using a dissipation inequality in integral form along with relaxing the requirement of smoothness of the Lyapunov function to mere continuity, which is done in [Ang02] and [Ang09] considering the incremental L^2-to-L^∞ (i.e., integral) versions of GAS and ISS for continuous-time systems, respectively.

More recently, time-discounted variants of i-IOSS were proposed in [KM20; ART21] for discrete-time systems, where it was shown that discounting past disturbances appears very natural and even without loss of generality. A corresponding converse Lyapunov result is provided in [ART21], which is structurally easier and more intuitive to establish with such a discount factor than without, as is the case in, e.g., [LSW96; KSW01; Ang02; Ang09]. Moreover, i-IOSS with time-discounting and its associated Lyapunov function are crucial for recent results in the field of optimization-based state estimation for discrete-time systems, compare [KM23; AR21; Sch+23].

1.2.2. Robust stability of MHE

One of the main concerns in MHE theory (and observer design in general) is to ensure, under appropriate conditions, that the corresponding estimation error is bounded and converges to zero in the ideal, unperturbed case, so that the unknown true trajectory can be recovered (at least asymptotically). To this end, an MHE scheme for continuous-time systems was proposed and analyzed in [MM95]. Since a cost function without a prior weighting was used (which can be seen as a regularization term), the system must satisfy an observability condition to ensure exponential convergence of the estimation error. Using such a cost function, however, requires long estimation horizons to ensure satisfactory performance in practice, compare [RMD20, Sec. 4.3.1]. Since the application of MHE inevitably requires some sort of sampling strategy (i.e., discrete time points at which the optimization is performed), schemes for discrete-time systems have recently been the main focus in the literature. Early results in the context of nonlinear systems employed certain uniform observability properties, compare, for example, [MR95; RRM03; ABB08].

In recent years, the notion of i-IOSS has proven to be a very useful concept for nonlinear detectability, enabling significant advances in MHE theory. In particular, in [RJ12], the authors established robust stability of FIE for i-IOSS systems, considering the special case of convergent (i.e., vanishing) disturbances. Robust stability of MHE in the more general and practically relevant case of persistent bounded disturbances was established in [Ji+16], albeit requiring a cost function that does not allow for standard least squares objectives. This was addressed in [Mül17] and generalized in [AR19b], however, yielding theoretical guarantees that—counter-intuitively—deteriorate with an increasing estimation horizon. The Lyapunov-based approach proposed in [AR21] is able to avoid this drawback, but on the other hand requires an additional stabilizability condition. Alternative approaches rely on an additional pre-estimating observer from which robust stability properties could be inherited [Liu13; GBE21].

In contrast, another line of research considers a cost function that includes explicit time discounting, which in the MHE context originates from the work [KM18]. This establishes a more direct link to the i-IOSS property and allows the derivation of strong robustness guarantees under less restrictive conditions, see, for example, [KM23; Hu24; Ale25]. In particular, the guarantees improve as the horizon length increases and do not require additional assumptions such as stabilizability or pre-estimating observers. The Lyapunov framework proposed in [Sch+23], which essentially relies on the same underlying principles, further simplifies the tuning and provides less conservative conditions on the horizon length sufficient for guaranteed robustly stable state estimation.

1.2.3. MHE for real-time applications

MHE requires solving a usually non-convex optimization problem at each time step, and is hence computationally demanding. Moreover, since the computing power available in practice is often severely limited, solving the optimization problem to global optimality at each time step is usually not possible within a fixed time interval.

In order to improve the real-time applicability of MHE, methods employing an additional *auxiliary observer* were developed to structurally simplify the optimization problem and thus save computing capacity. For example, in [SJF10], an MHE scheme for linear systems was proposed that utilized an additional Luenberger observer to replace the state equation as a dynamical constraint. As this allows to compensate for model uncertainties without computing an optimal disturbance sequence, the optimization variables could be reduced to one, namely the initial state at the beginning of the horizon. In [Suw+14], this idea was transferred to a class of nonlinear systems, and a major speed improvement compared to standard MHE could be shown. However, this results in a loss of degrees of freedom, since there is no possibility to tune the cost function with respect to model disturbances and measurement noise. In [Liu13], an observer was employed to construct a confidence region for the actual system state. Nevertheless, introducing this region as an additional constraint in the optimization problem can be quite restrictive and therefore may not allow significant improvements of MHE compared to the auxiliary observer. In [GBE21], a proximity-MHE scheme was proposed for a general class of nonlinear systems, where an additional observer is used to construct a stabilizing a priori estimate yielding a proper warm start for the low-level optimization algorithm, and nominal stability could be shown by Lyapunov arguments.

Nevertheless, all the above methods require optimal solutions to the (simplified, but still non-convex) MHE problem, and their complete computation within fixed time intervals is difficult (if not impossible) to guarantee. A more intuitive approach is to simply terminate the underlying optimization algorithm after a fixed number of iterations, which on the one hand provides only suboptimal estimates, but on the other hand ensures fixed computation times. However, since most results from the nonlinear MHE literature are crucially based on optimality [RMD20; Mül17; AR21; Hu24; KM23], stability of suboptimal MHE cannot be straightforwardly deduced. For practical (real-time) applications, it is therefore essential to develop suboptimal schemes that guarantee robust stability without requiring optimal solutions.

To this end, fast MHE methods were developed in [Küh+11; WVD14; AG17], performing only a predetermined number of iterations of a certain optimization algorithm (e.g., gradient- or Newton-based). However, the corresponding results rely on a strong uniform observability condition and (local) contraction properties of the specific algorithms, requiring both a proper initial guess and at least one iteration to ensure (local) stability, compare [WVD14; AG17]. In [WK17], the combination of a fast MHE scheme and pre-estimation using a nonlinear Luenberger observer was considered, combining the advantages of both approaches. A suboptimal proximity-MHE scheme for linear systems was proposed in [GGE22], where nominal stability guarantees could be given without performing any optimization by using a pre-stabilizing observer and contraction properties of a specific gradient-based optimization algorithm. This approach has recently been extended to nonlinear systems in [GGE21], thus providing nominal stability guarantees for a suboptimal nonlinear proximity-MHE scheme using local properties of the optimization algorithm involved. Whereas these algorithms require the computation of first-order sensitivities to perform the iterations, zero-order MHE methods were developed that completely avoid the online evaluation of sensitivities [BZD19] or use fixed approximations [Bau+21]. The resulting MHE schemes are suitable for real-time estimation of large-scale processes (arising, for example, from a discretization of partial differential equations), but their theoretical properties are of qualitative and local nature, and the respective conditions are hard to verify.

1.2.4. Joint state and parameter estimation

MHE is a model-based state estimation technique and hence requires knowledge of a suitable dynamical model of the system to be estimated. However, even if the general structure of the system is known, the model parameters are often uncertain and/or fluctuate during operation, e.g., due to heat production, mechanical wear, temperature changes or other external influences. This may invalidate the robustness guarantees, as they usually rely on an exact model of the real system and are therefore not necessarily valid in the case of parametric model uncertainties. In the worst case, this could even lead to the estimation error becoming unstable, compare, for example, [Fit71; SS71].

To address this problem, a min-max MHE scheme was proposed in [ABB12], where at each time step a least squares cost function is minimized for the worst case of the model uncertainties. However, such a min-max approach becomes computationally intensive for general nonlinear systems, and the worst-case consideration may be too conservative and affect estimation performance. In [MKZ23a], a regularization term was employed that depends on a given *a priori* estimate of the (constant) uncertain parameters, avoiding a nested min-max optimization scheme and ultimately yielding state estimates that are robust to changes in the true unknown parameter. Here, practical stability of the state estimation error with respect to the *a priori* parameter error could be established.

Yet it is often advantageous to not only ensure robustness against model errors, but also to obtain an estimate of the uncertain parameters, since a precise model is crucially required for, e.g., high-performance control, system monitoring, or fault detection. This demands suitable techniques for online parameter adaptation. In this context, an MHE scheme was proposed in [SJ11] by treating the unknown constant parameters as additional states with constant dynamics. The corresponding stability analysis is based on the transformation of the extended system into an observable and an unobservable but exponentially stable subsystem, where the temporary loss of observability (due to lack of excitation) is handled by suitable regularization and adaptive weights. However, the robustness properties have not been analyzed, and the imposed conditions for guaranteed state and parameter convergence are not trivial to verify in practice. In [FS23], MHE under a non-uniform observability condition is considered, which is potentially also suitable to be used for joint state and parameter estimation. The results, however, rely on persistently exciting inputs and, in particular, no fallback strategy is provided in case a lack of excitation occurs in practice during estimation. The work [BRD22] investigates MHE for joint state and parameter estimation from the perspective of numerical optimization. Here, the lack of excitation is addressed by using additional pseudo-measurements in case the variances of the estimates do not sufficiently decrease over the estimation horizon. This ensures that the corresponding covariance matrix remains bounded and the arrival cost is properly regularized; however, this approach lacks (global) stability guarantees.

An alternative approach to joint state and parameter estimation is provided by *adaptive observers*, which compute state estimates and simultaneously update internal model parameters. This concept originates from the work [Kre77] and has been extensively studied in the literature, see, e.g., [IS12] for an introduction to this topic. Theoretical guarantees usually consider the case of constant parameters and involve a detectability or observability condition on the system states and a persistence of excitation (PE) condition to establish parameter convergence. Different system classes (usually neglecting disturbances) have been considered, e.g., linear time-varying (LTV) systems [ŢB16], Lipschitz nonlinear systems under a linear parameterization [CR97], nonlinearly parameterized systems [Far+09; Tyu+13], or systems in a certain nonlinear adaptive observer canonical form, compare, e.g., [BG88; MST01]. An adaptive sliding mode observer was proposed in [EEZ16], which was generalized to a more general class of systems in [Fra+20], albeit under conditions that imply certain structural restrictions. Adaptive observers usually can also be applied to track (slowly) time-varying parameters if a forgetting factor is used in the design, see, e.g., [ŢB16]. Time-varying parameters are explicitly considered and analyzed in, e.g., [BG88] and [MST01], requiring that the parameter and its time-derivative are globally bounded for all times. Alternative approaches for systems in canonical forms can be found in, e.g., [BM21], where more general identifiers are used to estimate the dynamics.

The vast part of the literature on state and parameter estimation considers PE conditions to be uniform in time, which usually is restrictive and cannot be guaranteed *a priori* (except, e.g., for linear systems and suitable input trajectories). To ensure practical applicability, it is essential to investigate weaker, especially non-uniform, excitation conditions. In this context, for example, a regularized adaptive Kalman filter (for LTV systems) was proposed in [Mar+22] and an adaptive observer (for systems in a nonlinear adaptive observer canonical form) in [TM23]. In both works it could be shown that the state and parameter estimation errors are bounded without excitation and exponentially stable in the presence of PE. Relaxed excitation conditions have recently received much attention in the context of (pure) parameter estimation of regression models. In [EBO19], however, it was shown that weaker conditions than PE generally only allow for non-uniform asymptotic stability guarantees, which is also consistent with earlier works, e.g., [PLT01]. Using the *dynamic regressor extension and mixing* idea, exponential convergence could be established for linear regression models (and certain classes of nonlinear ones), merely assuming interval excitation (which is strictly weaker than uniform PE), compare, e.g., [Kor+22; ORA22].

1.2.5. Performance guarantees for state estimation

Current research in the field of MHE is primarily concerned with stability and robustness guarantees, see, for example, [RMD20, Ch. 4] and [ABB08; Ale+10; AR21; KM23; Sch+23; Hu24; Ale25]. These works essentially show that under suitable detectability conditions, the estimation error of MHE (i.e., the deviation between the estimated and the real system state) converges to a neighborhood of the origin, the size of which depends on the true unknown disturbance. However, results on the actual performance of nonlinear MHE methods, and in particular on the approximation accuracy and performance loss compared to a particular (challenging) benchmark, are lacking.

In general, a useful metric for quantifying the cumulative performance gap of a certain (estimation or control) algorithm with respect to a given benchmark is provided by the notion of dynamic regret. This is in fact a standard measure for analyzing related methods in the field of reinforcement learning [JOA10; ACJ21]. For the control of linear dynamical systems, regret-optimal controllers are designed in, e.g., [Sab+21; DSZ22; Mar+24b; Mar+24a]. Moreover, a regret analysis is performed for, e.g., online optimal control algorithms [Aga+19; LCL19; NM22], and the relation between bounded dynamic regret and asymptotic stability of the resulting closed loop is formally analyzed in [NM23].

In the context of state estimation for linear systems, regret-optimal filters are designed in [GH23; SH22], which essentially minimize the regret with respect to a clairvoyant (acausal) filter having access to future measurements. This approach is extended in [BDF23], where an exact solution to the minimal-regret observer is provided utilizing the system level synthesis framework. In [GGE22], an MHE scheme is proposed that provides regret guarantees with respect to an arbitrary comparative (e.g., the clairvoyant) observer. This approach is extended to nonlinear systems in [GGE21], but requires a restrictive convexity condition on the problem and disturbance- and noise-free data.

Whereas performance guarantees for state estimators are generally rather rare and usually restricted to linear systems, they often play an important role in nonlinear optimal control, especially when the overall goal is an economic one. Corresponding results usually employ a *turnpike* property of the underlying nonlinear optimal control problem, compare [McK86; CHL91]. This property essentially implies that optimal trajectories most of the time stay close to an optimal equilibrium (or in general an optimal time-varying reference), which is regarded as the *turnpike*. Turnpike-related arguments are an important tool for assessing the closed-loop performance of nonlinear model predictive controllers with general economic costs on finite and infinite horizons, see, for example, [Grü16; FGM18; GP19; FG22]. Necessary and sufficient conditions for the presence of the turnpike phenomenon in optimal control are discussed in, e.g., [Dam+14; GM16; Fau+22; Tré23], and are usually based on dissipativity, controllability, and suitable optimality conditions.

1.3. Contributions and outline of this thesis

The main contribution of this thesis is the development of MHE methods for general nonlinear systems in the presence of process disturbances and measurement noise, for which desired (and in particular not too conservative) robust stability and performance guarantees can be given under realistic and verifiable conditions. In the following, we outline the structure of this thesis and clarify the contributions in detail.

Chapter 2: Nonlinear detectability

In this chapter, we focus on i-IOSS as a characterization of detectability for general nonlinear systems. We start by introducing different notions of i-IOSS in discrete time in Section 2.1, encompassing the traditional asymptotic-gain formulation and modern, time-discounted versions. Then, in Section 2.2, we concentrate on continuous-time systems and propose a particular L^2-to-L^∞ variant of i-IOSS, namely time-discounted incremental integral IOSS (i-iIOSS). We introduce a corresponding Lyapunov function characterization of i-iIOSS relying on a dissipation inequality in integral form, where we show that an exponential decay can be considered without loss of generality. We establish equivalence between the existence of an i-iIOSS Lyapunov function and i-iIOSS by means of a converse Lyapunov theorem. Our proofs use similar tools as in previous works on incremental integral ISS [Ang09] and i-IOSS in the discrete-time setting [ART21]; however, we point out that the presented results do not straightforwardly follow from them. In particular, continuity of the Lyapunov function candidate is shown by replacing the standard local Lipschitz assumption on the vector field of the system by a global property involving the Osgood condition [Osg98]. As a byproduct, based on this assumption, we formally prove global existence and uniqueness of system trajectories by adapting the results from [Lip00; Bih56] to the generic class of measurable, locally essentially bounded functions.

Furthermore, we propose a time-discounted integral L^2-to-L^∞ variant of robust global asymptotic stability (RGAS) and show necessity of i-iIOSS for a system to admit a general observer mapping satisfying this property. Asking such a stability property from an observer is advantageous for several reasons: first, it can be seen as accounting for the disturbance energy under fading memory and thus allows for a physical interpretation; second, it directly implies an L^∞ error bound and thus combines the advantages of classical ISS and integral ISS properties. Overall, we provide a general framework for a Lyapunov-based robust stability analysis of observers in continuous time. This will be an essential tool in the context of moving horizon estimation in Chapter 3.

Chapter 3: Robust stability

In this chapter, we focus on robust stability guarantees for MHE and in particular concentrate on a recent Lyapunov-based MHE approach. We first provide a mathematical background on MHE by introducing a basic discrete-time MHE scheme in Section 3.1, where we discuss fundamental properties and characteristics. Then, we briefly introduce the Lyapunov-based MHE framework proposed in [Sch+23, Sec. III], which forms a basis for many of the results in this thesis (but is not itself a contribution[1] of it).

In Section 3.2, we propose a Lyapunov-based MHE scheme for general nonlinear continuous-time systems. We employ a least squares objective with fading memory and establish robust global exponential stability of the estimation error in a time-discounted L^2-to-L^∞ sense. Here, we heavily rely on the concepts of i-iIOSS and RGAS introduced in Chapter 2 to characterize the required detectability and robust stability properties. Our derivation builds on our ideas for the discrete-time case from [Sch+23, Sec. III]; however, the results do not trivially follow from this. Instead, the presented results are more general, require a different proof technique, and offer key advantages over purely discrete-time schemes, especially when the physical system to be estimated actually corresponds to a continuous-time one (which is often the case in practice). First, we note that arbitrary sampling strategies can be employed to define time instants at which the underlying optimization problem is actually solved, which can even be modified online at runtime. This provides a

[1] Julian D. Schiller (the author of this thesis) and Simon Muntwiler are joint first authors of the article [Sch+23]; Simon Muntwiler provided the theoretical analysis of discrete-time Lyapunov-based MHE (Sections III-B and III-C in [Sch+23]), which is part of the contributions of the PhD thesis [Mun24]; Julian D. Schiller contributed the comparison with existing results from the MHE literature (Section III-D), methods to verify the underlying detectability condition (Section IV), and the numerical examples (Section V), which are included in this thesis. A detailed description of the contributions of each author of [Sch+23] is given in Appendix A.