A moving horizon state and parameter estimation scheme with guaranteed robust convergence

Julian D. Schiller, Matthias A. Müller

Leibniz University Hannover, Institute of Automatic Control (e-mail: \{schiller,.mueller\}@irt.uni-hannover.de)

Abstract: We propose a moving horizon estimation scheme for joint state and parameter estimation for nonlinear uncertain discrete-time systems. We establish robust exponential convergence of the combined estimation error subject to process disturbances and measurement noise. We employ a joint incremental input/output-to-state stability (δ-IOSS) Lyapunov function to characterize nonlinear detectability for the states and (constant) parameters of the system. Sufficient conditions for the construction of a joint δ-IOSS Lyapunov function are provided for a special class of nonlinear systems using a persistence of excitation condition. The theoretical results are illustrated by a numerical example.

Keywords: Moving horizon estimation, state estimation, parameter estimation, parametric uncertainties, nonlinear systems, incremental system properties.

1. INTRODUCTION

Robust state estimation for nonlinear systems subject to noise is a problem of high practical relevance. However, if the underlying model is also uncertain and cannot accurately capture the real system behavior, the estimation error may even become unstable if this is not taken into account in the observer design. We address this issue by developing a robust moving horizon estimation (MHE) scheme for simultaneously estimating the states and (constant) parameters of a nonlinear system.

Recently, there have been significant developments in MHE theory, and in particular, strong robust stability properties have been established under practical conditions, see, e.g., (Allan and Rawlings, 2021; Knüfer and Müller, 2023; Schiller et al., 2022). Central to these results was the underlying detectability condition, namely δ-IOSS, which is necessary and sufficient for the existence of robustly stable state estimators, cf. (Allan et al., 2021; Knüfer and Müller, 2023). Procedures to verify this crucial property in practice have recently been developed in (Schiller et al., 2022). However, the derived guarantees for MHE rely on an exact model of the real system and are therefore not valid in the case of (parametric) model uncertainties. To address this problem, a min-max MHE scheme was proposed earlier in (Alessandri et al., 2012), 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. On the other hand, it is often beneficial not only to ensure robustness against model errors, but also to obtain an estimate of the uncertain parameters, since a good model is crucially required for, e.g., (high-performance) control, system monitoring, or fault detection. In this context, an MHE scheme was proposed in (Sui and Johansen, 2011), treating the 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 (in particular, the existence of a suitable transformation) for guaranteed state and parameter convergence are not trivial to verify in practice.

An alternative approach to simultaneous state and parameter estimation is provided by adaptive observers, the concepts of which have been extensively studied in the literature, see, e.g., the book by Ioannou and Sun (2012). Typically, this involves a detectability/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 systems (Guyader and Zhang, 2003), Lipschitz nonlinear systems under a linear parameterization (Cho and Rajamani, 1997), and nonlinearly parameterized systems (Farza et al., 2009). A more general class of nonlinear systems was considered in (Ibrir, 2018), however, requiring the verification of a large-dimensional linear matrix inequality (LMI), the feasibility of which is generally difficult to guarantee. Many results also consider a particular nonlinear adaptive observer canonical form (with dynamics affine in the unknown parameter and the nonlinearities depending on the measured output), cf., e.g., (Bastin and Gevers, 1988; Marino et al., 2001), where the latter also includes a robustness analysis, and (??), where lack of PE is countered. An adaptive sliding
mode observer was proposed in (Efimov et al., 2016), which was transferred to a more general class of systems in (Franco et al., 2020), albeit under conditions that imply certain structural restrictions. An alternative approach applicable to general nonlinear systems is provided by the supervisory framework, cf. e.g. Meijer et al. (2021), where, however, the desired degree of accuracy is directly related to the number of observers that have to be simulated and evaluated in parallel.

**Contribution** We propose an MHE scheme for joint state and parameter estimation of general nonlinear discrete-time systems subject to process disturbances and measurement noise. Our arguments are based on recent MHE results (Schiller et al., 2022; Allan and Rawlings, 2021), where only state (but no parameter) estimation is considered. In particular, we extend the concept of δ-IOSS by defining a joint δ-IOSS Lyapunov function for the system states and the (constant) parameters, which we then use as N-step Lyapunov function for MHE to establish robust convergence of the state and parameter estimates to their true values. As second contribution, we derive sufficient conditions for constructing such a joint δ-IOSS Lyapunov function for a certain class of nonlinear systems that are affine in the unknown parameter, which involves arguments from the adaptive observer literature, in particular, a PE condition similar to those from Guyader and Zhang (2003); Efimov et al. (2016); Ibrir (2018).

**Notation** The set of integers is denoted by \( \mathbb{I} \), the set of all integers greater than or equal to \( a \) for any \( a \in \mathbb{I} \) by \( \mathbb{I}_{\geq a} \), and the set of integers in the interval \([a, b]\) for any \( a, b \in \mathbb{I} \) with \( a \leq b \) by \( \mathbb{I}[a,b] \). The \( n \times n \) identity matrix is denoted by \( I_n \) and the \( n \times m \) matrix containing only zeros is denoted by \( 0_{n \times m} \), where we omit the indices if the dimension is unambiguous from the context. The weighted Euclidean norm of a vector \( x \in \mathbb{R}^n \) with respect to a positive definite matrix \( Q = Q^T \) is defined as \( \|x\|_Q = \sqrt{x^T Q x} \) with \( \|x\| = \|x\|_{I_n} \); the minimal and maximal eigenvalues of \( Q \) are denoted by \( \lambda_{\min}(Q) \) and \( \lambda_{\max}(Q) \), respectively. Given two matrices \( A = A^T \) and \( B = B^T \), we write \( A \geq B \) \( (A > B) \) if \( A - B \) is positive semi-definite (positive definite). For \( A, B \) positive definite, the maximum generalized eigenvalue (i.e., the largest scalar \( \lambda \) satisfying \( \lambda(A - \lambda B) = 0 \)) is denoted by \( \lambda_{\max}(A, B) \).

2. PROBLEM SETUP

We consider the nonlinear uncertain discrete-time system

\[
\begin{align*}
    x_{t+1} &= f_s(x_t, u_t, d_t, \theta), \quad (1a) \\
    y_t &= h_s(x_t, u_t, d_t, \theta), \quad (1b)
\end{align*}
\]

with state \( x_t \in \mathbb{X} \subseteq \mathbb{R}^n \), control input \( u_t \in \mathbb{U} \subseteq \mathbb{R}^m \), disturbances \( d_t \in \mathbb{D} \subseteq \mathbb{R}^p \), noisy output \( y_t \in \mathbb{Y} \subseteq \mathbb{R}^p \), and time \( t \in \mathbb{I}_{\geq 0} \). The nonlinear continuous functions \( f_s \) and \( h_s \) represent the system dynamics and the output equation, respectively, both of which depend on an unknown but constant parameter \( \theta \in \Theta \subseteq \mathbb{R}^p \).

The overall goal is to compute, at each time \( t \in \mathbb{I}_{\geq 0} \), the estimates \( \hat{x}_t \) and \( \hat{\theta}_t \) of the current state \( x_t \) and the real parameter \( \theta \). We provide conditions under which the combined estimation error

\[
e_t = \begin{bmatrix} e_{x,t} \\ e_{\theta,t} \end{bmatrix} \quad \text{with} \quad e_{x,t} = x_t - \hat{x}_t, \ e_{\theta,t} = \theta - \hat{\theta}_t \tag{2}
\]

robustly exponentially converges to a neighborhood around the origin, i.e., that there exist constants \( c_1, c_2 \geq 0 \) and \( \rho_1, \rho_2 \in [0,1) \) such that

\[
\|e_t\| \leq \max \left\{c_1 \rho_1^t \|e_0\|, \max_{j \in [1,t]} c_2 \rho_2^{t-j} \|d_{t-j}\| \right\} \tag{3}
\]

for all \( t \in \mathbb{I}_{\geq 0} \). Note that this directly implies \( e_t \to 0 \) if \( d_t \to 0 \) for \( t \to \infty \). To this end, a suitable detectability property of system (1) is required. As discussed in Section 1, the use of the δ-IOSS concept for this purpose has been central to recent advances in the field of nonlinear MHE for state estimation, (cf., e.g., (Schiller et al., 2022; Allan and Rawlings, 2021; Knüfer and Müller, 2023; Allan et al., 2021)). Therefore, it seems natural to extend this notion by considering both the states and the parameters of system trajectories via the following Lyapunov function definition.

**Definition 1.** (Joint δ-IOSS Lyapunov function). A function \( U : \mathbb{I}_{\geq 0} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^+ \) is a (quadratically bounded) δ-IOSS Lyapunov function for the system (1) on some set \( \mathcal{Z}_T \) if there exist a constant \( T \in \mathbb{I}_{\geq 1} \), matrices \( \mathcal{M}, \mathcal{M} > 0, Q, R \geq 0 \), and a contraction rate \( \lambda \in [0,1) \) such that for all trajectory pairs

\[
(x_t, u_t, d_t, \theta)_{t=0}^{K-1}, (\hat{x}_t, u_t, \hat{d}_t, \hat{\theta})_{t=0}^{K-1} \in \mathcal{Z}_T \tag{4}
\]

with \( x_{t+1} = f_s(x_t, u_t, d_t, \theta) \) and \( \hat{x}_{t+1} = f_s(\hat{x}_t, u_t, \hat{d}_t, \hat{\theta}) \) for all \( t \in [0,K-1] \) and any \( K \in \mathbb{I}_{\geq T} \), the following is satisfied:

\[
\begin{align*}
    \left\| x_t - \hat{x}_t \right\|_Q^2 &\leq U(t, x_t, \hat{x}_t, \theta, \hat{\theta}) \\
    \left\| \begin{bmatrix} x_t - \hat{x}_t \\ \theta - \hat{\theta} \end{bmatrix} \right\|_{\mathcal{M}}^2 &\leq U(t, x_t, \hat{x}_t, \theta, \hat{\theta}) \\
    t &\in [0,K], \tag{5a}
\end{align*}
\]

\[
\begin{align*}
    U(t+1, x_{t+1}, \hat{x}_{t+1}, \theta, \hat{\theta}) &\leq \lambda U(t, x_t, \hat{x}_t, \theta, \hat{\theta}) + \|d_t - \hat{d}_t\|_Q^2 + \|y_t - \hat{y}_t\|_R^2 \\
    t &\in [0,K-1], \tag{5b}
\end{align*}
\]

where \( y_t = h_s(x_t, u_t, d_t, \theta) \) and \( \hat{y}_t = h_s(\hat{x}_t, u_t, \hat{d}_t, \hat{\theta}) \).

It is intuitively clear that requiring the δ-IOSS Lyapunov function \( U \) to satisfy the conditions (5) for arbitrary system trajectories is not reasonable in practice. Therefore, we consider only certain pairs of trajectories of at least length \( T \) for which (5) is satisfied, contained in some set \( \mathcal{Z}_{T} \); later we will define \( \mathcal{Z}_T \) by means of a certain PE condition. Such a condition, the satisfaction of which usually requires the choice of a suitable input sequence \( \{u_{1:t}\}_{t=0}^{K-1} \), can then be mapped into the Lyapunov function due to the fact that we impose an additional explicit dependence of \( U \) on time, which is in contrast to the time-invariant definition typically used in the MHE literature (cf., e.g., (Schiller et al., 2022; Allan and Rawlings, 2021)). In the following, we use such a joint δ-IOSS Lyapunov function to design an MHE scheme for joint state and parameter estimation that guarantees robust exponential convergence of the combined estimation error (2) as characterized in (3).

3. MOVING HORIZON JOINT STATE AND PARAMETER ESTIMATION

At each time \( t \in \mathbb{I}_{\geq 0} \), the proposed MHE scheme considers the available input and output sequences of the system (1) within a moving horizon of length \( N_t = \min\{t, N\} \) for
some $N \in \mathbb{N}_0$. The current state and parameter estimates are obtained by solving the following nonlinear program:

$$
\min_{\hat{x}_{t-N_t}, u_j, d_j, \theta_j, d_0} J_t(\hat{x}_{t-N_t}, \hat{\theta}_t, d_0) \quad (6a)
$$

subject to

$$
\begin{align*}
\hat{x}_{j+1} &= f_s(\hat{x}_j, u_j, d_j, \hat{\theta}_t), \quad j \in \mathbb{N}_0, \\
\hat{\theta}_{t+1} &= h_s(\hat{x}_j, u_j, \hat{d}_j, \hat{\theta}_t), \quad j \in \mathbb{N}_0, \\
\hat{x}_j &= x_j, \quad j \in \mathbb{N}_0, \\
\hat{d}_j &= d_j, \quad j \in \mathbb{N}_0.
\end{align*}
$$

The decision variables $\hat{x}_{t-N_t}$ and $\hat{\theta}_t$ are obtained from solving the following nonlinear program:

$$
J_t(\hat{x}_{t-N_t}, \hat{\theta}_t, d_0) = 2N \left( \hat{x}_{t-N_t} - x_{t-N_t} \right)^T \Gamma \left( \hat{x}_{t-N_t} - x_{t-N_t} \right) + \sum_{j=1}^{N_t} \lambda_{j-1} \left( 2\|d_{j-1}\|_Q^2 + \|d_j - d_{j-1}\|_Q^2 \right),
$$

where $\hat{x}_{t-N_t}$ and $\hat{\theta}_t$ are the estimates obtained at time $t$, and $\{\hat{d}_{j-1}\}_{j=1}^{t-N_t}$ is the estimated output sequence of system (1), the parameters $\lambda, Q, R$ are from the $\delta$-JOSS Lyapunov function (cf. Definition 1), and $\Gamma$ is a constant weighting matrix specified below. We denote $\hat{x}_t$ as the minimizer of (7) and (6b) by $\hat{x}_{t-N_t}$, and the corresponding optimal state sequence by $\{\hat{x}_{j}\}_{j=1}^{t-N_t}$. The resulting state and parameter estimates at time $t \in \mathbb{N}_0$ are then given by $\hat{x}_t = \hat{x}_{t-N_t}$ and $\hat{\theta}_t = \hat{\theta}_{t-N_t}$, respectively. The following theorem provides an $N$-step Lyapunov function for the combined estimation error (2), from which robust exponential convergence can be directly deduced.

Theorem 2. Suppose that the system (1) admits a joint $\delta$-JOSS Lyapunov function $U$ according to Definition 1 on some $\mathbb{Z}_T$. Let $\Gamma = M$ in (7) and choose $N \in \mathbb{N}_0$ large enough such that $N \geq T$ and $\rho := \lambda_{\text{max}}(M, M) \lambda_N^2 < 1$. Suppose that for each $t \in \mathbb{N}_0$, the real and the currently estimated trajectory form a pair satisfying

$$
(x_j, u_j, d_j, \theta_j)_{j=1}^{t-N_t} \in \mathbb{Z}_T.
$$

Then, $W_N(x, \hat{x}, \hat{\theta}, \hat{\theta}) := U(N(x, \hat{x}, \hat{\theta}))$ is a joint $N$-step Lyapunov function on $\mathbb{Z}_T$ for the state and parameter estimation error $e_t$ (2) satisfying

$$
W_N(x_t, \hat{x}_t, \hat{\theta}_t) \leq \rho^N W_N(x_{t-N_t}, \hat{x}_{t-N_t}, \hat{\theta}_{t-N_t}) + 4 \sum_{j=1}^{N} \lambda_{j-1} \|d_{j-1}\|_Q^2
$$

for all $t \in \mathbb{N}_2N$.

Proof. Since (8) applies, we can evaluate the $\delta$-JOSS Lyapunov function $U$ along the pair formed by the real and the estimated trajectory on the estimation horizon. The rest of the proof follows by suitably adapting the arguments from the proof of (Schiller et al., 2022, Th. 1), where we exploit the relation between the specific structure (and parameterization) of the cost function (7) and the $\delta$-JOSS Lyapunov function $U$ satisfying (5).

Remark 3. (Convergence.) The joint $N$-step Lyapunov function (9) is only valid for all $t \in \mathbb{N}_{2N}$ due to the definition of $W_N$. However, if for $t \in \mathbb{N}_{0-N}$ the bounds (5) also hold for the pairs of trajectories of length $t < N$ (which generally is the case for causal—i.e., non-anticipating—$U$ via the initialization at time zero, compare the conditions of Theorem 10 below), then one can straightforwardly deduce that $e_2$ (2) satisfies (3) for all $t \in \mathbb{N}_0$ with $c_1 = 4 \sqrt{\lambda_{\text{max}}(M, M)}$, $c_2 = \frac{4}{1 - \rho} \sqrt{\lambda_{\text{max}}(M, M)}$, $\rho_1 = \sqrt{\rho}$, and $\rho_2 = \sqrt{\rho}$ by using similar arguments that were applied in the proofs of Theorem 2 and (Schiller et al., 2022, Cor. 1).

From a theoretical point of view, Theorem 2 is not surprising, as it is a fairly straightforward extension of recent results from the MHE context (in particular (Schiller et al., 2022), where only state (but no parameter) estimation was considered). The underlying problem is rather the construction of $\delta$-JOSS Lyapunov functions satisfying Definition 1; choosing $U$ quadratic with respect to a constant positive definite weighting matrix, for example, which would be a first naive attempt, is in general not feasible in (5b). In the following section, we provide an alternative approach to the construction of a time-varying quadratic $\delta$-JOSS Lyapunov function for a particular class of nonlinear systems using arguments from the adaptive observer literature.

### 4. SUFFICIENT CONDITIONS FOR THE CONSTRUCTION OF A JOINT $\delta$-JOSS LYAPUNOV FUNCTION

Throughout the following, we restrict ourselves to the special case where $f$ is affine in $\theta$ and subject to additive disturbances $d$ and with $h$ linear, that is, we consider

$$
\begin{align*}
\Phi_{s}(x, u) &= f_s(x, u) + G(x, u)\theta + Ed_t, \\
h_s(x, u, \theta) &= Cx_t + Fd_t
\end{align*}
$$

with constant matrices $E, C, F$ of appropriate dimensions and with $G$ having the following properties.

Assumption 4. (Boundedness.) The function $G$ is continuously differentiable and the sets $X, U, \Theta$ are compact.

Let us consider two trajectories of system (1) given by $(x_t, u_t, d_t, \theta_t)_{t=0}^{K-1}$ and $(\tilde{x}_t, \tilde{u}_t, \tilde{d}_t, \tilde{\theta}_t)_{t=0}^{K-1}$ for any $K \in \mathbb{N}_0$. Using the mean value theorem, we first note that

$$
\begin{align*}
f(x, u) - f(\tilde{x}, u) &= A(x, \tilde{x}, u)(x - \tilde{x}) \quad (11)
\end{align*}
$$

with

$$
\begin{align*}
A(x, \tilde{x}, u) := \int_0^1 \frac{\partial f}{\partial x}(\tilde{x} + s(x - \tilde{x}), u)ds
\end{align*}
$$

for all $x, \tilde{x} \in X$ and $u \in U$.

Assumption 5. (Detectability.) There exists a continuous mapping $L : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^p$ such that the matrix

$$
\Phi(x, \tilde{x}, u) := A(x, \tilde{x}, u) + L(x, \tilde{x}, u)C
$$

satisfies

$$
\begin{align*}
\Phi(x, \tilde{x}, u)^T P \Phi(x, \tilde{x}, u) \leq \mu P
\end{align*}
$$

for some $P > 0$ and $\mu \in [0, 1)$ uniformly for all $x, \tilde{x} \in X$ and $u \in U$. 

Remark 6. (Computation of $P, L$). The matrices $P$ and $L$ can be easily computed by transforming (14) into an LMI condition using standard linear algebra and a suitable verification (e.g., using sum-of-squares (SOS) optimization). The additional degree of freedom provided by allowing $L$ to depend on both the states $x$ and $\bar{x}$ can be used, e.g., to compensate for nonlinear terms in (12), which is beneficial in several aspects, compare Remark 11. We point out that this is generally not possible in the context of (adaptive) observers due to the fact that $L$ is crucially required for performing the observer recursions at each time step and therefore enforced to be constant as in, e.g., (Ibiris, 2018).

Having set the assumptions regarding the system state, we also need assumptions regarding the parameters. To this end, note that for any fixed $\theta$, the definition $G_\theta(x, u) := G(x, u, \theta)$ represents a continuously differentiable function $G_\theta^T : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$. Therefore, by the mean value theorem,

$$(G(x, u) - G(\bar{x}, u))\theta = G'_\theta(x, u) - G'_\theta(\bar{x}, u) = G_\theta(x, u)(x - \bar{x}),$$

where

$$G_\theta(x, \bar{x}, u) := \int_0^1 \frac{\partial G'_\theta(x, s(x - \bar{x}), u)}{\partial x} ds$$

for all $x, \bar{x} \in \mathbb{X}$ and $u \in \mathbb{U}$.

Assumption 7. There exists a matrix $H$ such that $G_\theta(x, \bar{x}, u)^T P G_\theta(x, \bar{x}, u) \preceq C^T H C$ uniformly for all $x, \bar{x} \in \mathbb{X}$, $u \in \mathbb{U}$, and $\theta \in \Theta$ with $P$ from Assumption 5.

Remark 8. (Conditions on $G$). Condition (17) is linear in $H$ and thus can be easily verified using standard LMI methods. We point out that Assumption 7 is related to the condition used in (Cho and Rajamani, 1997); furthermore, $H$ always exists for the special case where $G(x, u) = G(Cx, u)$, which includes the important forms of nonlinear adaptive observer canonical forms that are often considered in the adaptive observer literature (cf., e.g., (?Marino et al., 2001)).

We define the recursions

$$Y_{t+1} = \Phi(x_t, \bar{x}_t, u_t)Y_t + G(\bar{x}_t, u_t), \quad t \in [0, K-1]$$

$$S_{t+1} = \eta_S + Y_{t}^T C^T C Y_t + \epsilon \quad \in [0, K-1]$$

for some fixed $Y_0 \in \mathbb{R}^{n \times o}$, $S_0 \in \mathbb{R}^{o \times o}$ and $\eta \in (0, 1)$. As usual in the context of system identification/parameter estimation, we need a PE condition in order to guarantee exponential convergence of the parameter estimation error, which in our case characterizes the set $Z_T$ from Definition 1. More precisely, for some fixed $T \in \mathbb{Z}_{\geq 0}, \alpha > 0$, we define

$$Z_T := \left\{ \left( x_{t-1}, u_{t-1}, \theta_{t-1} \right)^K_{t=0}, (\bar{x}_{t}, u_{\bar{t}}, \bar{\theta}_{t})_{t=0}^{K-1} \right\}$$

$$\in \left( \mathbb{X}^K \times \mathbb{U}^K \times \mathbb{D}^K \times \Theta^K \times \Theta \right)^T : K \in \mathbb{Z} \geq T,$$

$$x_{t+1} = f(x_t, u_t, \theta_t), \quad t \in [0, K-1],$$

$$\bar{x}_{t+1} = \bar{f}(x_t, \bar{x}_t, u_t, \bar{\theta}_t), \quad t \in [0, K-1],$$

$$\sum_{j=1}^T Y_{t-j}^T C^T C Y_{t-j} \geq \alpha I_o, \quad t \in [T, K].$$

Our PE condition is essentially similar to those appearing in the adaptive observer literature (cf., e.g., (Guyader and Zhang, 2003; Efimov et al., 2016; Ibiris, 2018)), albeit we enforce it by considering only trajectory pairs belonging to $Z_T$. Note that the a priori verification for the general case is still an open problem; instead, membership of trajectory pairs to the set $Z_T$ can be verified by simulations, or—in special cases—order, compare Remark 11. A trajectory pair in $Z_T$ directly implies positive definiteness and uniform boundedness of $S_t$ as the following lemma shows.

Lemma 9. Let some $Y_0$ and $S_0 > \alpha I_o$ be given. Then, there exist $\sigma_1, \sigma_2 > 0$ such that

$$\sigma_1 I_o \preceq S_t \preceq \sigma_2 I_o$$

for all $t \in [0, K]$ and all trajectory pairs satisfying (4) with $Z_T$ from (20) for any $K \in \mathbb{Z} \geq T$.

The proof is shifted to the appendix. We are now in the position to state the following result.

Theorem 10. Let Assumptions 4, 5, and 7 hold and let

$$U(t, x_t, \bar{x}_t, \theta, \tilde{\theta}) := \left\| \left[ x_t - \bar{x}_t \right] - \left[ \theta - \tilde{\theta} \right] \right\|_{M_t}^2, \quad t \in [0, K]$$

with

$$M_t := \begin{bmatrix} P & -P Y_t \bar{Y}_t^T P \bar{Y}_t \bar{P} + a S_{\tilde{t}} \end{bmatrix} = I \quad \in [0, K].$$

Then, there exists some $a > 0$ small enough such that $U$ (22) is a joint $\delta$-ISS Lyapunov function on $Z_T$ according to Definition 1, i.e., (5) is satisfied for all trajectory pairs satisfying (4) with $Z_T$ from (20) for any $K \in \mathbb{Z} \geq T$.

Proof. Inspired by adaptive observer literature (compare, e.g., (Guyader and Zhang, 2003; Bastin and Gevers, 1988)), we employ the (filtered) transformation

$$z_t = T_t(z_t), \quad \tilde{z}_t = T_t(\tilde{z}_t), \quad T_t = \begin{bmatrix} I - Y_t & 0 \end{bmatrix}.$$ (24)

Note that the time-varying transformation matrix $T_t$ is non-singular independent of $t$ and $K$. Using (23) and (24), observe that $U$ in (22) can now be equivalently re-written as

$$U(t, x_t, \bar{x}_t, \theta, \tilde{\theta}) := \left\| \left| z_t - \tilde{z}_t \right| + a \check{\theta} - \check{\theta} \right\|_{S_t}^2, \quad t \in [0, K].$$ (25)

Define $W(z_t, \tilde{z}_t) := \left\| z_t - \tilde{z}_t \right\|_2^2$, and $V(t, \theta, \tilde{\theta}) = \left\| \theta - \tilde{\theta} \right\|_2^2$. The remaining proof is structured in three parts: we first establish Lyapunov-like properties for $W$, second for $V$, and finally for $U$.

Part 1. By the dynamics (1a) with (10a), the difference $z_t - \tilde{z}_t$ evolves according to

$$z_{t+1} - \tilde{z}_{t+1} = x_{t+1} - \bar{x}_{t+1} - Y_{t+1}(\theta - \tilde{\theta})$$

$$\begin{split}
\leq f(x_t, u_t) - f(\bar{x}_t, u_t) - \Phi(x_t, \bar{x}_t, u_t)Y_t(\theta - \tilde{\theta}) + D_G(x_t, \bar{x}_t, u_t, \theta) + D_E(d_t, \tilde{d}_t),
\end{split}$$

where $D_G(x_t, \bar{x}_t, u_t, \theta) := (G(x_t, u_t) - G(\bar{x}_t, u_t))$. Adding $0 = L(x_t, \bar{x}_t, u_t)(y_t - \bar{y}_t - (y_t - \bar{y}_t))$ with $y_t = C x_t + \bar{d}_t$ and $\bar{y}_t = C \bar{x}_t + F \bar{d}_t$ by (1b) and (10b) to the right-hand side together with application of (11) and (13) yields

$$z_{t+1} - \tilde{z}_{t+1} = \Phi(x_t, \bar{x}_t, u_t)(z_t - \tilde{z}_t) + (E + L(x_t, \bar{x}_t, u_t)F)(d_t - \tilde{d}_t) + D_G(x_t, \bar{x}_t, u_t, \theta) - L(x_t, \bar{x}_t, u_t)(y_t - \bar{y}_t).$$

Using the Cauchy-Schwarz inequality together with fact that $(a_1 + a_2)^2 \leq (1 + \epsilon)a_1 + (1/\epsilon)a_2$ for any $\epsilon > 0, a_1, a_2 \geq 0$, we obtain that

$$z_{t+1} - \tilde{z}_{t+1} \leq (1 + \epsilon)\alpha I_o + (1/\epsilon)\alpha I_o.$$
where

By Assumption 5, we note that

By using (15), (17), and (10b). Application of (27) and (28) were applied to obtain (26) together with the definition boundedness of $S(0_{\overline{0}}, 1)$ in (31) and define $(\lambda, Q, R)$ such that $4\kappa V_N < 1$ for some fixed $\kappa > 0$, so that the cost function (7) can be implemented given suitable values for $\lambda, Q, R$. Now for any $t \in I_{\overline{0}, N}$, let $Y_{j|t}, S_{j|t}, M_{j|t}$ with $j \in I_{[0, N]}$ represent $Y_j, S_j, M_j$ given the currently estimated trajectory $(\hat{x}^t_{t-N_{t-j}+j} \mid u_{t-N_{t-j}+j}, \hat{d}_{t-N_{t-j}+j}^t, \hat{\theta}_{t-N_{t-j}+j}^t)$, i.e., determined by (18), (19), and (23), respectively, with $\hat{x}_j := \hat{x}^t_{t-N_{t-j}+j}$. Note that $Y_j(t)$ (and hence $S_j(t)$ and $M_j(t)$) can explicitly be computed, since the recursion rule (18) only depends on known quantities in case of a constant $\Phi$ (in particular, it does not depend on the unknown true state system $x_j$). Hence, for a given value of $\alpha > 0$, we can verify the following two conditions online:

\[
\lambda_{\text{max}}(M_0, M_{N_{t-j}}) \leq \kappa, \quad \forall t \in I_{\overline{0}, N}. \tag{33}
\]

Satisfaction of (34) directly implies that condition (8) of Theorem 2 holds with $T = N$. If also (33) holds, we can replace $\lambda_{\text{max}}(\overline{M}, M)$ by $\kappa$ in the proof of Theorem 2 using the facts that $\Gamma = M_0$ and $U$ (22) is quadratic. Consequently, if the conditions (33) and (34) are satisfied and $\lambda, Q, R$ are chosen appropriately, then the estimation error $e_t$ (2) is guaranteed to robustly exponentially converge as characterized in (3), compare Remark 3 and see also the simulation example in Section 5.

5. NUMERICAL EXAMPLE

To illustrate our results, we consider the following system

\[
x_1^t = x_1 + t \Delta b_1 (x_2 - a_1 x_1 - a_2 x_1^2 + a_3 x_1^3) + d_1,
\]
\[
x_2^t = x_2 + t \Delta (x_1 - x_2 + x_3),
\]
\[
x_3^t = x_3 - t \Delta x_2 + d_3,
\]
\[
y = x_1 + d_4,
\]

which corresponds to the Euler-discretized modified Chua's circuit system from Yang and Zhao (2015) using the step size $t_s = 0.01$ under additional disturbances $d \in \mathbb{R}^3$, where only $x_1$ can be measured. The parameters are $b_1 = 12.8, b_2 = 19.1, a_1 = 0.6, a_2 = -1.1, a_3 = 0.45$, which leads to a chaotic behavior of the system. We consider the initial condition $x_0 = [2, 0.1, -2]^\top$, assume that $x$ evolves in the
choose the parameters \( a, \epsilon \) constant, leading to the contraction rate \( \mu \) (L " ofberg, 2009) in combination with the semidefinite pro-

7 are verified in Matlab using the SOS toolbox Yalmip

 contained in the set \( \Theta = [0 \times 1 \times [0,1] \times [-3,3] \), and choose the naive initial estimate \( \hat{x}_0 = [0,0,0]^T \). Furthermore, we assume that the exact parameter \( a_3 = \theta \) is unknown but contained in the set \( \Theta = [0,2,0,8] \). In the following, we treat \( d \) as a uniformly distributed random variable with \( |d_i| \leq 10^{-2}, i = 1,2,3 \) for the process disturbance and \( |d_1| \leq 0.1 \) for the measurement noise. Assumptions 5 and 7 are verified in Matlab using the SOS toolbox Yalmip (L"ofberg, 2009) in combination with the semidefinite programming solver Mosek (MOSEK ApS, 2019). Here, we choose \( L \) linear in \( x \) and \( \hat{x} \) such that \( \phi \) in (13) becomes constant, leading to the contraction rate \( \mu = 0.9 \), and we choose the parameters \( a, \epsilon_1, \epsilon_2 \) such that \( \hat{\mu} = \eta = \lambda = 0.911 \). Then, we verify Assumption 7 on \( X \) and compute the matrices \( Q, R \). Since \( \phi \) is constant, we can apply Remark 11, where we choose \( N = 200 \) such that \( 4\kappa \lambda^N < 1 \) with \( \kappa = 10^7 \). Figure 1 shows the estimation error of each state and the parameter as well as the combined estimation error (2) normalized to unity over time, revealing exponen-
tial convergence to a neighborhood around the origin as guaranteed by Theorem 2 (using the modifications from Remark 11 and the fact that conditions (33) and (34) are satisfied during the simulation). The relatively large parameter estimation error in the time interval [400,600] is due to temporary weak excitation, which comes into play especially in combination with the measurement noise present here. The estimation results could be easily im-
proved by choosing \( N \) larger or by temporarily stopping parameter estimation when weak excitation is detected.

6. CONCLUSION

We proposed a moving horizon estimation scheme for joint state and parameter estimation and provided sufficient conditions for guaranteed robust exponential convergence. In particular, we used a joint \( \delta \)-IOSS Lyapunov function as a detectability condition for the system states and parameters to construct an \( N \)-step Lyapunov function for the combined state and parameter estimation error. Moreover, we provided a systematic approach for the construction of \( \delta \)-IOSS Lyapunov functions for the special class of nonlinear systems that are affine in the parameters using a persistence of excitation condition. Extensions to treat more general classes of nonlinear systems and to relax some of the conditions are the subject of future work.

REFERENCES

Alessandri, A., Baglietto, M., and Battistelli, G. (2012). Min-max moving-horizon estimation for uncertain discrete-time linear systems. SIAM J. Control Optim., 50(3), 1439–1465.

Allan, D.A. and Rawlings, J.B. (2021). Robust stability of full information estimation. SIAM J. Control Optim., 59(5), 3472–3497.

Allan, D.A., Rawlings, J.B., and Teel, A.R. (2021). Nonlinear detectabil-

ity and incremental input/output-to-state stability. SIAM J. Control Optim., 59(4), 3017–3039.

Bastin, G. and Gevers, M. (1996). Stable adaptive observers for nonlinear time-varying systems. IEEE Trans. Autom. Control, 33(7), 650–658.

Cho, Y.M. and Rajamani, R. (1997). A systematic approach to adaptive observer synthesis for nonlinear systems. IEEE Trans. Autom. Control, 42(4), 534–537.

Efimov, D., Edwards, C., and Zolghadri, A. (2016). Enhancement of adaptive observer robustness applying sliding mode techniques. Automatica, 72, 53–56.

Farza, M., M’Saaid, M., Maatoug, T., and Kamoun, M. (2009). Adaptive observers for nonlinearly parameterized class of nonlinear systems. Automatica, 45(10), 2292–2299.

Franco, R., Rios, H., Efimov, D., and Perruquetti, W. (2020). Adaptive estimation for uncertain nonlinear systems with measurement noise: A sliding-mode observer approach. Int. J. Robust Nonlinear Control, 31(9), 3809–3826.

Guyader, A. and Zhang, Q. (2003). Adaptive observer for discrete time linear time varying systems. IFAC Proceedings Volumes, 36(16), 1705–1710.

Ibrir, S. (2018). Joint state and parameter estimation of non-linearly parameterized discrete-time nonlinear systems. Automatica, 97, 226–233.

Ioannou, P. and Sun, J. (2012). Robust Adaptive Control. Dover Publications, Inc., Mineola, NY, USA.

Kn"{u}fer, S. and M"{u}ller, M.A. (2023). Nonlinear full information and moving horizon estimation: Robust global asymptotic stability. Automatica, 150, 110603.

L"ofberg, J. (2009). Pre- and post-processing sum-of-squares programs in practice. IEEE Trans. Autom. Control, 54(5), 1007–1011.

Marino, R., Santosusso, G., and Tomei, P. (2001). Robust adaptive observers for nonlinear systems with bounded disturbances. IEEE Trans. Autom. Control, 45(6), 967–972.

Meijer, T., Dolk, V., Chong, M., Postoyan, R., de Jager, B., Nesic, D., and Heemels, W. (2021). Joint parameter and state estimation of noisy discrete-time nonlinear systems: A supervisory multi-observer approach. In 60th IEEE Conf. Decision Control. IEEE.

MOSEK ApS (2019). The MOSEK optimization toolbox for MATLAB manual. Version 9.0. [Online]. Available: http://docs.mosek.com/9.0/toolbox/index.html.

Schiller, J.D., Muntwiler, S., K"{o}hler, J., Zeilinger, M.N., and M"{u}ller, M.A. (2022). A Lyapunov function for robust stability of moving horizon estimation. arXiv:2302.15744.

Sui, D. and Johansen, T.A. (2011). Moving horizon observer with regularisation for detectable systems without persistence of excitation. Int. J. Control, 84(6), 1041–1054.

Yang, J. and Zhao, L. (2015). Bifurcation analysis and chaos control of the modified chua’s circuit system. Chaos Solit. Fractals, 77, 332–339.

Appendix A. PROOF OF LEMMA 9

Proof. We start by noting that the recursion (19) implies \( S_t = \eta^t S_0 + \sum_{j=1}^{t} \eta^{t-j} \Gamma_{t-j} C^T Y_{t-j} \), \( t \in \mathbb{N} \), which directly yields the lower bound \( S_t \geq \eta^{-t} \alpha L_0 \) for \( t \in \mathbb{N} \).

For the upper bound, we obtain that \( \| S_t \| \leq \eta^t \| S_0 \| + \sum_{j=1}^{t} \eta^{t-j} \| C Y_{t-j} \| \) for \( t \in \mathbb{N} \). Since \( Y_t \) satisfies (18), we note that \( Y_t \) is (uniformly) upper bounded for all \( t \in \mathbb{N} \) due to the fact that \( \Phi \) satisfies (14) and \( G \) is uniformly bounded using Assumption 4. Hence, there exists some \( \beta > 0 \) such that \( \| C Y_t \| \leq \beta \) uniformly for all \( t \in \mathbb{N} \). Finally, by applying the geometric series and defining \( \sigma_1 := \eta^{-1} \alpha \) and \( \sigma_2 := \left( \| S_0 \| + \mathbb{E}[\beta^2] \right) \), we derive (21), which concludes this proof.