Generic Stability Implication From Full Information Estimation to Moving-Horizon Estimation

Wuhua Hu

Abstract—Optimization-based state estimation is useful for handling of constrained linear or nonlinear dynamical systems. It has an ideal form, known as full information estimation (FIE), which uses all past measurements to perform state estimation, and also a practical counterpart, known as moving-horizon estimation (MHE), which uses most recent measurements of a limited length to perform the estimation. This work reveals a generic link from robust stability of FIE to that of MHE, showing that the former implies at least a weaker robust stability of MHE, which implements a longer enough horizon. The implication strengthens to strict robust stability of MHE if the corresponding FIE satisfies a mild Lipschitz continuity condition. The revealed implications are then applied to derive new sufficient conditions for robust stability of MHE, which further reveals an intrinsic relation between the existence of a robustly stable FIE/MHE and the system being incrementally input/output-to-state stable.

Index Terms—Disturbances, full information estimation (FIE), incremental input/output-to-state stability (i-IOSS), moving-horizon estimation (MHE), nonlinear systems, robust stability, state estimation.

I. INTRODUCTION

Optimization-based state estimation is an estimation approach that performs state estimation by solving an optimization problem. Compared with conventional approaches, such as Kalman filtering (KF), which deals with linear dynamical systems, and its extensions, such as extended KF and unscented KF, which can deal with nonlinear dynamical systems based on linearization techniques, optimization-based approach has the advantage of handling linear and nonlinear dynamical systems directly and also including various physical or operational constraints [1]. The optimization formulation also admits the flexible definition of the objective function to be optimized, which in some cases is necessary to accurately recover the state [2].

Optimization-based state estimation generally takes one of the two forms: full information estimation (FIE), which uses all past measurements to perform the estimation, and moving-horizon estimation (MHE), which uses most recent measurements of a limited length to perform the estimation. MHE is a practical approximation of FIE, which is ideal and computationally intractable. The interest in FIE lies in two aspects: serving as a benchmark to MHE and providing useful insights for the stability analysis of MHE. Recent studies on FIE and MHE are concentrated on the analyses of their stability and robustness, as are critical to guide the applications of MHE. While earlier literature assumes restrictive and idealistic conditions, such as observability and/or zero or a priori known convergent disturbances [3], [4], [5], [6], [7], [8], [9], recent literature considers more practical conditions, such as detectability and/or in the presence of bounded disturbances [1], [10], [11].

An important progress was made in [11], which introduced the incremental input/output-to-state stability (i-IOSS) concept on the detectability of nonlinear systems (as developed in [12]) to study robust stability of FIE and MHE. Given an i-IOSS system, it was shown that under mild conditions, FIE is robustly stable and convergent for convergent disturbances, whereas conditions for FIE and MHE to be robustly stable under bounded disturbances were posted as an open research challenge. Hu et al. [13] provided a prompt response to this challenge, identifying a general set of conditions for FIE to be robustly stable for i-IOSS systems. Here onward, a series of research works have been inspired to close the challenge.

A particular type of cost functions with a max-term was investigated for a class of i-IOSS systems in [14], establishing robust stability of the FIE. The conditions were enhanced in [15], enabling robust stability also of MHE if a sufficiently long horizon is applied. The conclusion was extended to MHE without a max-term in the cost function [16]. Meanwhile, it was shown that MHE is convergent for convergent disturbances, with or without a max-term. On the other hand, Hu [2] revealed an implication link from robust stability of FIE to that of MHE, and consequently identified rather general conditions for MHE to be robustly stable by inheriting conditions which ensure robust stability of the corresponding FIE. By making use of a Lipschitz continuity condition introduced in [2], Allan and Rawlings [17] streamlined and generalized the analysis and results of [16], showing that the key is essentially to assume that the system satisfies a form of exponential detectability. When global exponential detectability is assumed, the MHE can further be shown to be robustly globally exponentially stable (RGES) by implementing properly time-discounted stage costs [18].

The reviewed robust stabilities of FIE and MHE were concluded based on the concept defined in [11], which is, however, found to be flawed in that such defined robustly stable estimator does not imply convergence for convergent disturbances [19], [20]. This motivates a necessary modification of the stability concept in [19], which redefines the estimate error bound with a worst-case time-discounted instead of uniformly weighted impact of the disturbances. The new concept is an enhancement of the old one, and was shown to be compatible with the original i-IOSS detectability of a system, which is necessary for establishing robust stability of both FIE and MHE. With this new concept, it becomes straightforward to understand the earlier robust local/global exponential stability results reported of FIE and MHE [16], [18]. It also motivates some new results as developed in [1, Ch. 4], which introduced a kind of stabilizability condition to establish robust stability of FIE and MHE. Despite conceptual elegance, the new condition involves an inequality, which is uneasy to verify in general.

Motivated by the new and stronger concept of robust stability, this work aims to motivate general sufficient conditions for robust stability

1The main results were obtained in the late of 2012, although the paper was not able to be accepted for publication until 2015.
of MHE by first establishing a generic implication link from robust stability of FIE to that of MHE, and then transforming the challenge into identifying sufficient conditions for ensuring robust stability of FIE. While the reasoning approach is inspired by the ideas introduced in [2], the contributions of this work are twofold.

1) A generic implication is established from robust stability of FIE to practical robust stability of the corresponding MHE, and the implication becomes stronger to robust stability of the MHE if the FIE admits a certain Lipschitz continuity property.

2) Given the implication, new sufficient conditions are derived for robust stability of MHE by first developing those for the corresponding FIE. An interesting finding is that a system being i-IOSS is necessary but also sufficient for the existence of a robustly stable FIE. Consequently, it is a similar but weaker conclusion applicable to MHE.

In particular, due to space limit, this work focuses on presenting global stability results while their local counterparts are left to a separate report, as referred to [21].

The rest of this article is organized as follows. Section II introduces notation, setup, and necessary preliminaries. Section III defines general forms of FIE and MHE, and introduces robust stability concepts. Section IV reveals the implication from global robust stability of FIE to that of the corresponding MHE. Section V applies the implication to establish robust stability of MHE, by first developing conditions for ensuring robust stability of the corresponding FIE. Finally, Section VI concludes this article.

II. NOTATION, SETUP, AND PRELIMINARIES

The notation mostly follows the convention in [1] and [2]. The symbols \( \mathbb{R} \), \( \mathbb{R}_{\geq 0} \), and \( \mathbb{I}_{\geq 0} \) denote the sets of real numbers, nonnegative real numbers, and nonnegative integers, respectively, and \( \mathbb{I}_{a:b} \) denotes the set of integers from \( a \) to \( b \). The constraints \( t \geq 0 \) and \( t \in \mathbb{I}_{\geq 0} \) are used interchangeably to refer to the set of discrete times. The symbol \( \cdot \) denotes the Euclidean norm of a vector. The bold vector \( \mathbf{x}_{\mathbb{I}_{a:b}} \) denotes a sequence of vector-valued variables \( (x_a, x_{a+1}, \ldots, x_b) \), and with a function \( f \) acting on a vector \( x \), \( f(\mathbf{x}_{\mathbb{I}_{a:b}}) \) stands for the sequence of function values \( f(x_a), f(x_{a+1}), \ldots, f(x_b) \). Given a scalar function \( f : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \), any \( t \in \mathbb{R} \) and any scalar function \( g \), the notation \( f^m(g(t)) \) refers to the \( m \)-fold composition of function \( g \) by function \( f \) subject to the given second argument \( t \), and \( f^0 \) equals the identity function.

Throughout this article, \( t \) refers to a discrete time, and as a subscript, it indicates dependence on time \( t \), whereas the superscripts or subscripts \( x, w, \) and \( e \) are used exclusively to indicate a function or variable that is associated with the state \( x \), process disturbance \( w \), or measurement noise \( e \). The symbols \( \lfloor x \rfloor \) and \( \lceil x \rceil \) refer to the integers that are closest to \( x \) under and above, respectively. Given two scalars \( a \) and \( b \), let \( a \oplus b := \max(a, b) \). The operator \( \oplus \) is both associative and commutative, i.e., \( (a \oplus b) \oplus c = a \oplus (b \oplus c) \) and \( a \oplus b = b \oplus a \), and, furthermore, is distributive with respect to (w.r.t.) increasing functions.

That is, \( \alpha(a \oplus b \oplus c) = \alpha(a) \oplus \alpha(b) \oplus \alpha(c) \) if the function \( \alpha \) is increasing in the argument. The frequently used \( K, L \), and \( KL \) functions are defined as follows.

**Definition 1 (K, L, and KL functions):** A function \( \alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} \) is a \( K \) function if it is continuous, zero at zero, and strictly increasing. A function \( \varphi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} \) is a \( L \) function if it is continuous, nonincreasing, and satisfies \( \varphi(t) \to 0 \) as \( t \to \infty \). A function \( \beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0} \) is a \( KL \) function if, for each \( t \geq 0 \), \( \beta(\cdot, t) \) is a \( K \) function and for each \( s \geq 0 \), \( \beta(s, \cdot) \) is a \( L \) function.

Consider a discrete-time system described by

\[
x_{t+1} = f(x_t, u_t), \quad y_t = h(x_t) + v_t \tag{1}
\]

where \( x_t \in X \subseteq \mathbb{R}^n \) is the system state, \( w_t \in W \subseteq \mathbb{R}^q \) is the process disturbance, \( y_t \in Y \subseteq \mathbb{R}^p \) is the measurement, and \( v_t \in V \subseteq \mathbb{R}^r \) is the measurement disturbance, all at time \( t \). Here, we study state estimation as an independent subject, and so control inputs (if there were any) known up to the estimation time are treated as given constants, which do not cause difficulty to later defined optimization and related analyses and, hence, are neglected in the problem formulation for brevity [1], [11]. The functions \( f \) and \( h \) are assumed to be continuous and known.

Let \( \bar{x}_0 \) be an a priori estimate of the initial state \( x_0 \). The following uncertainties setup is assumed.

**Assumption 1:** The uncertainty \( x_0 - \bar{x}_0 \) is bounded for all \( x_0, \bar{x}_0 \in X \), and so are the uncertainties \( (w_t, v_t) \) for all \( t \geq 0 \).

Consider two state trajectories. Let

\[
\begin{align*}
\pi_0 &:= x^{(1)}_0 - \bar{x}_0, \\
\pi_{\tau+1} &:= w^{(1)}_\tau - w^{(2)}_\tau, \quad \pi_{\tau+1} := h(x^{(2)}_{\tau+1}) - h(x^{(1)}_{\tau+1})
\end{align*}
\tag{2}
\]

for all \( \tau \in \mathbb{I}_{0:t-1} \), where \( \pi_{\tau+1} \) corresponds to \( v^{(1)}_\tau - v^{(2)}_\tau \) if additive measurement noises are present while identical measurements are assumed. Hence, \( \pi_{0:2t} \) collects a sequence of deviation vectors, and its domain is denoted as \( \Pi_t \). Let \( \iota(\tau) \) extract the time index (i.e., the abovementioned original index \( \tau \) of \( \pi \)). Specifically, we have

\[
\iota(\pi_{\tau+1}) = \iota(\pi_{\tau+1}) = \tau, \quad \forall \tau \in \mathbb{I}_{0:t-1}.
\tag{3}
\]

Note that the time index of \( \pi_0 \) is defined as \( -1 \), which refers to the time associated with a priori information.

With the abovementioned notation, we can have an equivalent statement of the i-IOSS definition given in [1] and [19], to facilitate later presentation.

**Definition 2 (i-IOSS):** The system \( x_{t+1} = f(x_t, u_t), y_t = h(x_t) \) is i-IOSS if there exists \( \alpha \in KL \) such that for every two initial states \( x^{(1)}_0 \) and \( x^{(2)}_0 \), and two sequences of disturbances \( w^{(1)}_{0:t-1} \) and \( w^{(2)}_{0:t-1} \), the following inequality holds for all \( t \in \mathbb{I}_{0:t} \):

\[
\|x^{(1)}_t - x^{(2)}_t\|_2 \leq \max_{i \in \mathbb{I}_{0:t-1}} \alpha(\|\pi_i\|, t - \iota(\pi_i) - 1) \tag{4}
\]

where \( x^{(i)}_t \) is a shorthand of \( x_i(x^{(i)}_0, w^{(i)}_{0:t-1}) \) for \( i \in \{1, 2\} \). Furthermore, the system is exponentially i-IOSS (exp-i-IOSS) if \( \alpha \) admits an exponential form as \( \alpha(s, \tau) := c e^{\lambda s} \) with certain \( \lambda \in (0, 1), c > 0 \) and for all \( s, \tau \geq 0 \).

It is straightforward to show that the abovementioned definition is equivalent to the one introduced in [19], which applies KL functions \( \alpha_e, \alpha_w, \) and \( \alpha_v \) instead of \( \alpha \) for \( i = 0 \), all \( i \in \mathbb{I}_t \) and all \( i \in \mathbb{I}_{0:t-1} \), respectively, in (4), and as proved in [19], such definition is equivalent to the original i-IOSS definition introduced in [12], which applies \( \alpha \) instead of KL function to each uncertainty \( |\pi_i| \) for all \( i \geq 1 \).

The following defines Lipschitz continuity of a function at a given point, as to be used in later analysis.

**Definition 3 (Lipschitz continuity at a point):** Given a subset \( S \subseteq \mathbb{R}^n \), a function \( f : \mathbb{R}^n \to \mathbb{R}^m \) is said to be Lipschitz continuous at a point \( x^* \in S \) over the subset \( S \) if there is a constant \( c \) such that \( |f(x) - f(x^*)| \leq c|x - x^*| \) for all \( x \in S \).

III. OPTIMIZATION-BASED STATE ESTIMATION

Consider the system described in (1). Given a present time \( t \), the state estimation problem is to find an optimal estimate of state \( x_t \) based on historical measurements \( \{y_s\} \) for all \( \tau \) in a time set. Ideally, all measurements up to time \( t \) are used, leading to the so-called FIE; practically, only measurements are used within a limited distance backward from time \( t \), yielding the so-called MHE. Both FIE and MHE
can be cast as optimization problems.\(^2\) To be concise, we will first define MHE and then treat FIE as a variant.

Let MHE implement a moving horizon of size \(T\). The decision variables are denoted as \((\chi_{T-\tau}, \omega_{T-\tau}, \nu_{T-\tau})\), which correspond to the system variables \((\chi_{T-\tau}, \omega_{T-\tau}, \nu_{T-\tau})\), and let the optimal decision variables be \((\hat{\chi}_{T-\tau}, \hat{\omega}_{T-\tau}, \hat{\nu}_{T-\tau})\). Since \(\hat{\chi}_{T-\tau}\) is uniquely determined from \(\hat{\chi}_{T-\tau}\) and \(\hat{\omega}_{T-\tau}\), the decision variables essentially reduce to \((\hat{\chi}_{T-\tau}, \hat{\omega}_{T-\tau})\).

In addition, let \(\hat{\chi}_{T-\tau}\) be a priori estimate of \(\chi_{T-\tau}\). Without loss of generality, the prior estimates of the disturbances are assumed to be zero. Denote the cost function as \(V_T^0(\chi_{T-\tau}, \hat{\chi}_{T-\tau}, \hat{\omega}_{T-\tau}, \hat{\nu}_{T-\tau})\), which penalizes uncertainties in the initial state, the process, and the measurements. Then, the MHE instance at time \(t\) is defined by the following optimization problem:

**MHE (or FIE)**

\[
\text{if } T \leftarrow t : \quad \min_{0} V_{\chi}(\chi_{T-\tau} - \hat{\chi}_{T-\tau}, \hat{\chi}_{T-\tau})
\]

s.t. \(\chi_{T-\tau} = f(\chi_{T-\tau}, \nu_{T-\tau}) \forall r \in I_{T-\tau} \forall t\)

\[
y_{T-\tau} = h(\chi_{T-\tau}) + \nu_{T-\tau} \forall r \in I_{T-\tau} \forall t\)

\[
\chi_{T-\tau} \in \mathcal{X}, \chi_{T-\tau} \in \mathcal{W}, \chi_{T-\tau} \in \mathcal{V},
\]

As \(\nu_{T-\tau} \in \mathcal{Y}\) is uniquely determined by \(\chi_{T-\tau} \) and \(\omega_{T-\tau}\), it is kept mainly for the convenience of expressing the disturbance set and the objective function. Since the global optimal solution \(\hat{\chi}_{T-\tau}\) for any \(\tau \leq t\), is dependent on time \(t\) when the MHE instance is defined, to be unambiguous we use \(\hat{\chi}_{T-\tau}^\tau\) to represent \(\hat{\chi}_{T-\tau}\) that is solved from the instance defined at time \(t\). This keeps \(\hat{\chi}_{T-\tau}^\tau\) unchanged, while the realization \(\hat{\chi}_{T-\tau}\) varies as MHE renews itself in time.

To define FIE, it suffices to adapt the horizon size \(T\) in the MHE formulation to taking the time-varying value \(t\). The yielded FIE has the form of an MHE, but with complete data originating from the zero initial time. To link them easily, an FIE is called the corresponding FIE of an MHE based on which the FIE is derived, and conversely is instancewise unique.

**Definition 4 (Robust stable estimation):** The estimate \(\hat{\chi}_{T-\tau}\) of state \(x_{T-\tau}\) is based on partial or full sequence of the noisy measurements, \(y_{0:T} = h(x_{0:T}, \omega_{0:T-\tau}) + \nu_{0:T}\). The estimate is robustly globally asymptotically stable (RGAS) if there exist functions \(\beta_{\nu}, \beta_{w}\), and \(\beta_{o} \in K\mathcal{L}\) such that the following inequality holds for all \(x_{0:T}, \hat{x}_{0:T-\tau} \in \mathcal{W}, \nu_{0:T-\tau} \in \mathcal{V}, \) and \(t \in I_{0:T}\):

\[
|x_{T-\tau} - \hat{x}_{T-\tau}| \leq \beta_{\nu}(|x_{0:T} - \hat{x}_{0:T}|, t)
\]

\[
\oplus \max_{\nu \in \mathcal{W}} \max_{\nu_{T-\tau}} \beta_{\nu}(|*\nu|, t - \tau - 1).
\]

If, furthermore, the \(\mathcal{K}\mathcal{L}\) function admits an exponential form as \(\beta_{\nu}(s, \tau) := e^{c_{\nu} s}\) with certain \(c_{\nu} > 0\), and for all \(s \in \{x, w, v\}\), then the estimate is said to be RGES.

\(^2\)Readers are referred to \([11]\) for a brief introduction of their connection to control problems and difference from probabilistic formulations.

\(^3\)As in \([1]\), the last measurement is not considered for ease of presentation, although the inclusion does not change the conclusions.

The last measurement \(y_{T}\) and, hence, the corresponding noise \(v_{T}\) are not considered in the abovementioned inequality, to keep the definition consistent with the formulations of FIE and MHE. Here, the definition of RGAS strengthens the one introduced in \([11]\), which applies \(K\mathcal{L}\) instead of \(K\mathcal{C}\) functions to the disturbances. This change is necessary to enable a desirable feature that a state estimator, which is RGAS, must be convergent under convergent disturbances \([11, 19]\). The next definition presents a weaker alternative of the abovementioned robust stability, which will also be needed in later analysis.

**Definition 5 (Practical robust stable estimation):** The estimate defined in Definition 4 is practically RGAS (pRGAS) if given any \(\epsilon > 0\), there exist functions \(\beta_{\nu}, \beta_{w}\), and \(\beta_{o} \in K\mathcal{L}\) such that the following inequality holds for all \(x_{0:T}, \hat{x}_{0:T-\tau} \in \mathcal{W}, \nu_{0:T-\tau} \in \mathcal{V}, \) and \(t \in I_{0:T}\):

\[
|x_{T-\tau} - \hat{x}_{T-\tau}| \leq \epsilon \oplus \beta_{\nu}(|x_{0:T} - \hat{x}_{0:T}|, t)
\]

\[
\oplus \max_{\nu \in \mathcal{W}} \max_{\nu_{T-\tau}} \beta_{\nu}(|*\nu|, t - \tau - 1).
\]

The adverb “practically” before RGAS is employed to keep it in line with the practical stability concept developed in control literature (e.g., \([22]\)). Compared with an RGAS estimate, the pRGAS estimate admits a lower bound with a nonvanishing constant term \(\epsilon\). As will be shown later, this term can be made arbitrarily small if MHE implements a long enough horizon.

**IV. STABILITY IMPLICATION FROM FIE TO MHE**

At any discrete time, an MHE instance can be interpreted as the corresponding FIE initiating from the start of the horizon over which the MHE instance is defined. Thus, the corresponding FIE being robustly stable implies that each MHE instance is robustly within the time horizon over which the instance is defined. If we interpret this as MHE being instancewise robustly stable, then the challenge reduces to identifying conditions under which instancewise robust stability implies robust stability of MHE. This observation was made in \([2]\), and the challenge was solved there for a weaker definition of RGAS. This section resolves the challenge subject to the stronger stability concept given by Definition 4.

To that end, as in \([2]\), we apply an ordinary assumption on the prior estimate \(\hat{x}_{T-\tau}\) of the initial state \(x_{T-\tau}\) of an MHE instance.

**Assumption 2:** Given any time \(T \geq T + 1\), the prior estimate \(\hat{x}_{T-\tau}\) of \(x_{T-\tau}\) is given such that

\[
|x_{T-\tau} - \hat{x}_{T-\tau}| \leq |x_{T-\tau} - \hat{x}_{T-\tau}|.
\]

The assumption is obviously true if \(\hat{x}_{T-\tau}\) is set to \(\hat{x}_{T-\tau}\) itself, which is the MHE estimate obtained at time \(t - T\). Alternatively, a better \(\hat{x}_{T-\tau}\) might be obtained with smoothing techniques, which use measurements both before and after time \(t - T\) \([11, 23]\).

Next, we present an important lemma that links global robust stability of MHE with that of its corresponding FIE.

**Lemma 1 (Stability implication from FIE to MHE):** If FIE is RGAS as per \((6)\), then there exists \(T \in I_{0:T}\) such that the corresponding MHE under Assumptions 1 and 2 is pRGAS for all \(T \geq T\). If, further, the \(K\mathcal{L}\) function \(\beta_{\nu}(\cdot, \cdot, \cdot)\) in \((6)\) is globally Lipschitz continuous at the origin, then the implication strengthens to the existence of \(T \in I_{0:T}\) such that the MHE is RGAS for all \(T \geq T\) under Assumption 2. If, furthermore, the \(K\mathcal{L}\) functions \(\beta_{w}(\cdot, \cdot, \cdot)\) and \(\beta_{\nu}(\cdot, \cdot, \cdot)\) are also globally Lipschitz continuous at the origin, then the MHE is RGES for all \(T \geq T\) under Assumption 2.

**Proof:** RGAS implies pRGAS. Let \(n := \frac{1}{T}\), and so \(0 \leq t - nT \leq T - 1\). For all \(r \in I_{0:T-1}\), the MHE and the corresponding FIE estimates are the same, both denoted as \(\hat{x}_{T-\tau}^\tau\). So, given any \(T \in I_{0:T-1}\), the absolute
estimation error $|x_\tau - \hat{x}_\tau|$ satisfies the RGAS inequality given by (6).
That is, we have

$$
| x_\tau - \hat{x}_\tau | \leq \beta_x ( (x_\tau - \bar{x}_0), \tau )
\quad \oplus \quad \max_{e \in \{ x, v \}} \max_{t \in \mathbb{I}_{1:n} - 1} \beta_x ( |x_t - \hat{x}_t|, t - \tau - 1 )
$$

(8)

for all $\tau \in \mathbb{I}_{0, T-n-1}$. Next, we proceed to show that the RGAS property is maintained for all $\tau \in \mathbb{I}_{-n+1,T}$. Let $t_1 := 0$ and $t_i := -(n - i) T$ for all $i \in \mathbb{I}_{0, n}$. The MHE instance at time $t_1$ can be viewed as the corresponding FIE confined to time interval $[\bar{t}_0, t_1]$. Thus, the MHE satisfies the RGAS property within this interval. That is, by (6), we have the following:

$$
| x_{t_1} - \hat{x}_{t_1} | \leq \beta_x ( (x_{t_1} - \bar{x}_{t_1}), T )
\quad \oplus \quad \max_{e \in \{ x, v \}} \max_{t \in \mathbb{I}_{1:n} - 1} \beta_x ( |x_t - \hat{x}_t|, t_1 - \tau - 1 )
$$

where Assumption 2 has been applied to produce the first term of the right-hand side of the inequality. Repeat this reasoning for the MHE instance defined at time $t_2$ and then apply the abovementioned inequality, yielding

$$
| x_{t_2} - \hat{x}_{t_2} | \leq \beta_x ( (x_{t_1} - \bar{x}_{t_1}), T )
\quad \oplus \quad \max_{e \in \{ x, v \}} \max_{t \in \mathbb{I}_{1:n} - 1} \beta_x ( |x_t - \hat{x}_t|, t_2 - t_1 - \tau - 1 )
\leq \beta_x^2 ( (x_0 - \bar{x}_0), T )
\quad \oplus \quad \max_{e \in \{ x, v \}} \max_{t \in \mathbb{I}_{1:n} - 1} \beta_x ( |x_t - \hat{x}_t|, t_1 - t_2 - \tau - 1 )
\quad \oplus \quad \max_{e \in \{ x, v \}} \max_{t \in \mathbb{I}_{1:n} - 1} \beta_x ( |x_t - \hat{x}_t|, t_2 - t_1 - \tau - 1 )
$$

By induction, we obtain

$$
| x_{t \tau} - \hat{x}_{t \tau} | = | x_{t_{n-1}} - \hat{x}_{t_{n-1}} | \leq |x_0 - \bar{x}_0| + \sum_{l=1}^{n-1} \max_{e \in \{ x, v \}} \max_{t \in \mathbb{I}_{1:n} - 1} \beta_x ( |x_t - \hat{x}_t|, t_{l+1} - t_l - \tau - 1 )
$$

Since $|x_{t_0} - \hat{x}_{t_0}|$ satisfies inequality (8), this implies that

$$
| x_{t_\tau} - \hat{x}_{t_\tau} | \leq \beta_x^{n} ( \beta_x ( |x_0 - \bar{x}_0|, t_0 ), T )
\quad \oplus \quad \max_{e \in \{ x, v \}} \max_{t \in \mathbb{I}_{1:n} - 1} \beta_x^{n} ( |x_t - \hat{x}_t|, t_{l+1} - t_l - \tau - 1 ), T )
\quad \oplus \quad \max_{e \in \{ x, v \}} \max_{t \in \mathbb{I}_{1:n} - 1} \beta_x^{n} ( |x_t - \hat{x}_t|, t_2 - t_1 - \tau - 1 )
$$

for all $T \geq T_\tau$.

Moreover, we have

$$
| x_{t_\tau} - \hat{x}_{t_\tau} | \leq \beta_x^{n} ( \beta_x ( |x_0 - \bar{x}_0|, t_0 ), T )
\quad \oplus \quad \max_{e \in \{ x, v \}} \max_{t \in \mathbb{I}_{1:n} - 1} \beta_x^{n} ( |x_t - \hat{x}_t|, t_{l+1} - t_l - \tau - 1 )
\quad \oplus \quad \max_{e \in \{ x, v \}} \max_{t \in \mathbb{I}_{1:n} - 1} \beta_x^{n} ( |x_t - \hat{x}_t|, t_2 - t_1 - \tau - 1 )
$$

where the relation that $| x_{t_\tau} - \hat{x}_{t_\tau} | \leq \beta_x^{n} ( \beta_x ( |x_0 - \bar{x}_0|, t_0 ), T )
\quad \oplus \quad \max_{e \in \{ x, v \}} \max_{t \in \mathbb{I}_{1:n} - 1} \beta_x^{n} ( |x_t - \hat{x}_t|, t_{l+1} - t_l - \tau - 1 )
\quad \oplus \quad \max_{e \in \{ x, v \}} \max_{t \in \mathbb{I}_{1:n} - 1} \beta_x^{n} ( |x_t - \hat{x}_t|, t_2 - t_1 - \tau - 1 )$ (9)

for all $T \geq T_\tau$, where $\beta_x \in \mathcal{K}$ for all $e \in \{ x, v \}$ and $\beta_x ( |x_0 - \bar{x}_0|, t_0 )$. This implies that the MHE is pRGAS by definition, and, hence, completes the proof of the first conclusion.

**RGAS \Rightarrow pRGAS:** If further, $\beta_x ( \cdot, 0 )$ from (6) is globally Lipschitz continuous at the origin, then for any $T \geq 0$ there must exist a function $\mu_\beta$ such that $\beta_x ( s, \tau ) \leq \mu_\beta(\tau)$ for all $s \in \mathbb{X}$ because $\beta_x \in \mathcal{L}$. By applying this property to (9), for all $T \geq T_\tau$ satisfying $\mu_\beta(\mathbb{I}_0, t)$ with certain $\eta \in (0, 1)$, the inequality there proceeds as

$$
| x_\tau - \hat{x}_\tau | \leq \beta_x^2 ( \beta_x ( |x_0 - \bar{x}_0|, T )
\quad \oplus \quad \max_{e \in \{ x, v \}} \max_{t \in \mathbb{I}_{1:n} - 1} \beta_x ( |x_t - \hat{x}_t|, t - \tau - 1 )
$$

for all $T \geq T_\tau$. This implies that the MHE is pRGAS by definition, and, hence, completes the proof of the first conclusion.
\[
\mu_x \left[ \frac{\alpha}{2} \right] (T) \beta_x \left( \left| \pi \right| \alpha, t - \left\lceil \frac{t - \tau}{T} \right\rceil \right) T - \tau - 1.
\]

\[
\leq \eta^\frac{\alpha}{2} - 1 \mu_x (0) |x_0 - \bar{x}_0|
\]

\[
\oplus \max_{e \in [w, v]} \max_{t \in [0, t - 1]} \eta^\frac{\alpha}{2} - 1 \beta_x (|\pi|, 0). \quad (10)
\]

Consequently, the MHE is RGAS by definition, which completes the proof of the second conclusion.

**RGAS \Rightarrow RGES:** If in addition to \( \beta_x (\cdot, 0) \), the \( \mathcal{K} \) functions \( \beta_x (\cdot, 0) \) and \( \beta_x (\cdot, 0) \) from (6) are globally Lipschitz continuous at the origin, then inequality (10) proceeds as

\[
|x_t - \hat{x}_t^*| \leq \eta^\frac{\alpha}{2} - 1 \mu_x (0) |x_0 - \bar{x}_0|
\]

\[
\oplus \max_{e \in [w, v]} \max_{t \in [0, t - 1]} \eta^\frac{\alpha}{2} - 1 \mu_x (0) |\pi|
\]

where \( \mu_x (0) \) is the Lipschitz constant of \( \beta_x (\cdot, 0) \) at the origin for all \( s \in \{ w, v \} \). Consequently, the MHE is RGES by definition, which completes the proof.

**Lemma 1** indicates that RGAS of FIE implies pRGAS of the corresponding MHE which implements a long enough horizon, and that the implication strengthens to RGAS or RGES of the MHE if the FIE additionally satisfies certain global Lipschitz continuity conditions. Explicit computation of a valid horizon size is possible in both cases, as referred to [21] (this article is an extended version of this paper). It should also be noted that Assumption 1 on bounded uncertainties is used to establish the former case, which is nonetheless not required for the latter case. In the absence of Assumption 1, a valid horizon size will depend on realized magnitudes of the uncertainties, and consequently, only a semiglobal counterpart of the pRGAS will be established in the former case.

**Remark 1:** In the second case of Lemma 1, an FIE being RGAS and its bound function \( \beta_x (\cdot, 0) \) being globally Lipschitz at the origin imply that the system is exp-i-IOSS, which can be verified by extending the proof for the local case in [17, Prop. 2] to the global case. However, the two conditions do not necessarily imply that the FIE will be RGES (or at least the proof is unclear yet). On the other hand, if an FIE is RGES, then the Lipschitz conditions in the third case are valid and, hence, the corresponding MHE will be RGES by Lemma 1, as is in line with the existing results reported in [18] and [24].

**Remark 2:** Kneuer and Mueller [25] proved the existence of a finite horizon for MHE to be RGAS if the corresponding FIE induces a global contraction for the estimation error, which essentially requires the dynamical system to be exp-i-IOSS. As referred to Remark 1, this conveys the same necessary condition as in one particular case of Lemma 1.

### V. Robust Stability of FIE and MHE

This section presents new sufficient conditions for robust stability of FIE, and also for that of the corresponding MHE by applying the stability implication revealed in the last section.

To make the conditions easy to interpret and the proof concise to present, the following notations are introduced in the spirit of (2):

\[
\bar{x}_t := x_0 - \bar{x}_0, \pi_{t+1} := w_{t+1}, \pi_{t+1} := v_t
\]

\[
\bar{v}_t := v_0 - \bar{x}_0, \bar{v}_{t+1} := v_t - \bar{w}_t, \tilde{v}_{t+1} := v_t - \bar{v}_t
\]

\[
\bar{s}_t := x_0 - \bar{x}_0, \bar{s}_{t+1} := s_t - \bar{x}_0, \tilde{s}_{t+1} := s_t - \bar{s}_t
\]

for all \( \tau \in I_{0:t-1} \), where \( \bar{x} \) and \( \bar{v} \) refer to the optimal estimates of \( x \) and \( \pi \), respectively. Given any \( s \in \{ \pi, \bar{\pi}, \bar{x}, \tilde{s} \} \), notation \( s_{0:2t} \) collects the sequence of vector variables \( (\pi, t)_{0:2t} \) and the corresponding domain is denoted by \( \Pi_s \), and \( t(s) \) extracts the time index (i.e., the original index \( \tau \)) of \( s \) as per (3). Given any \( i \in I_{20:2t} \), it is easy to verify that \( t(s_i) = t(\bar{s}_i) = t(\tilde{s}_i) \) and that \( s_i = \bar{s}_i + \tilde{s}_i \).

The next two assumptions are introduced to establish robust stability of FIE.

**Assumption 3:** There exist \( \rho, \bar{\rho} \in K \mathcal{L} \) such that the cost function of FIE \( V_t (\bar{\pi}_{0:2t}) \), which is continuous, satisfies the following inequality for all \( \bar{\pi}_{0:2t} \in \Pi_s \) and \( t \in I_{20:2t} \):

\[
\max_{e \in [w, v]} \rho (|\pi_i|, t - t(\bar{s}_i) - 1) \leq V_t (\bar{\pi}_{0:2t})
\]

\[
\leq \max_{e \in [w, v]} \rho (|\pi_i|, t - t(\bar{s}_i) - 1). \quad (12)
\]

**Assumption 4:** The \( \mathcal{K} \) function \( \alpha \) in the i-IOSS property (4) and the \( \mathcal{K} \) functions \( \rho \) and \( \bar{\rho} \) in Assumption 3 satisfy the next inequality for all \( \bar{\pi}_{0:2t} \in \Pi_s \), \( \tau, \tau' \in I_{0:t+1} \), and \( t \in I_{20:2t} \):

\[
\alpha (2\bar{\rho}^{-1} (\rho (|\pi_i|, \tau), \tau'), \tau') \leq \bar{\rho} (|\pi_i|, \tau). \quad (13)
\]

for certain \( \bar{\rho} \in \mathcal{K} \), in which \( \rho^{-1} (\cdot, \cdot) \) is the inverse of \( \rho (\cdot, \cdot) \) w.r.t. its first argument given the second argument \( \tau \).

Overall, Assumption 3 requires that the FIE has a property that mimics the i-IOSS property of the system, whereas Assumption 4 ensures that the FIE is more sensitive than the system to the uncertainties so that accurate inference of the state is possible. The interpretation of the relatively more obscure Assumption 4 becomes clear with a concrete realization in the following.

**Lemma 2 (Concrete realization of Assumption 4):** Assumption 4 is true if the \( \mathcal{K} \) function \( \rho \) satisfies

\[
\rho (|\tilde{s}_i|, \tau) \geq \alpha (2|\tilde{s}_i|, \tau) \quad (14)
\]

for all \( \bar{\pi}_{0:2t} \in \Pi_s, \tau \in I_{0:t+1} \), and \( t \in I_{20:2t} \).

**Proof:** It suffices to show that inequality (13) is satisfied. Subject to (14), we have

\[
\alpha (2\rho^{-1} (\rho (s, \tau), \tau'), \tau') \leq \rho (\rho^{-1} (\rho (s, \tau), \tau'), \tau') = \rho (s, \tau)
\]

for all \( s, \tau, \tau' \) in applicable domains. This implies that (13) is satisfied with \( \bar{\rho} := \rho \) and hence completes the proof.

Note that Assumption 4 and Lemma 2, each add a condition to the bound functions of Assumption 3 but are both independent of the assumption. The robust stability of FIE/MHE can then be established under Assumptions 3 and 4.

**Theorem 1 (RGAS of FIE and pRGAS/RGES of MHE):** If the system is i-IOSS and the cost function of FIE satisfies Assumptions 3 and 4, then the FIE is RGAS and its corresponding MHE under additional Assumptions 1 and 2 with a sufficiently long horizon is pRGAS. If, in addition, the \( \mathcal{K} \) functions \( \alpha (\cdot, 0) \) and \( \bar{\rho} (\cdot, 0) \) in Assumption 4 are globally Lipschitz continuous at the origin, then the MHE is RGES.

**Proof:** The global optimal solution of \( \bar{\pi}_{0:2t} \) for the FIE is denoted as \( \bar{\pi}_{0:2t} \) [cf., (11)], yielding a minimum cost \( V_t \). It follows that for all \( t \geq 0 \)

\[
V_t = V_t (\bar{\pi}_{0:2t}) \leq V_t (\bar{\pi}_{0:2t}) \leq \max_{e \in [w, v]} V_t (\bar{\pi}_{0:2t}) \leq V_t.
\]

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Consequently, by Assumption 3, we have \( \rho(|\hat{\pi}_i|, t - \iota(\hat{\pi}_i) - 1) \leq V_i^0 \leq V_i \), and further \(|\hat{\pi}_i| \leq 2^{-1}(V_i, t - \iota(\hat{\pi}_i) - 1)\).

Since \( \hat{\pi}_i = \pi_i - \hat{\pi}_i \) for each \( i \in \mathbb{I}_{0:2t} \), by applying the triangle inequality this implies that

\[
|\hat{\pi}_i| \leq |\pi_i| + |\hat{\pi}_i| + 2^{-1}(V_i, t - \iota(\hat{\pi}_i) - 1) \\
\leq 2|\pi_i| + 2^2^{-1}(V_i, t - \iota(\hat{\pi}_i) - 1) \\
\leq 2|\pi_i| + 2^2^{-1} (\bar{V}_i, t - \iota(\hat{\pi}_i) - 1) \\
\leq 2|\pi_i| + 2^2^{-1} (\bar{V}_i, t - \iota(\hat{\pi}_i) - 1)
\]

for all \( i \in \mathbb{I}_{0:2t} \). Substituting (15) into the i-IOSS property of (4) yields

\[
|x(t, x_0, w_{0;1-1}) - x(t, \bar{x}_0, \bar{w}_{0;1-1})| \\
\leq \max_{i \in \mathbb{I}_{0:2t}} \alpha(|\pi_i|, t - \iota(\hat{\pi}_i) - 1) \\
\leq \max_{i \in \mathbb{I}_{0:2t}} \alpha(2|\pi_i|, t - \iota(\hat{\pi}_i) - 1) \\
\leq \max_{i \in \mathbb{I}_{0:2t}} \alpha(2|\pi_i|, t - \iota(\hat{\pi}_i) - 1) \\
\leq \max_{i \in \mathbb{I}_{0:2t}} \alpha(2|\pi_i|, t - \iota(\hat{\pi}_i) - 1) \\
\leq \max_{i \in \mathbb{I}_{0:2t}} \alpha(2|\pi_i|, t - \iota(\hat{\pi}_i) - 1) \\
\leq \max_{i \in \mathbb{I}_{0:2t}} \alpha(2|\pi_i|, t - \iota(\hat{\pi}_i) - 1) \\
\leq \max_{i \in \mathbb{I}_{0:2t}} \alpha(2|\pi_i|, t - \iota(\hat{\pi}_i) - 1) \\
\leq \max_{i \in \mathbb{I}_{0:2t}} \alpha(2|\pi_i|, t - \iota(\hat{\pi}_i) - 1)
\]

where the equality \( \iota(\hat{\pi}) = \iota(\pi) = \iota(\hat{\pi}) \) and Assumption 4 have been used to derive the last inequality, and \( \bar{\rho}(\pi_i, \tau) := \alpha(2|\pi_i|, \tau) \rho(\pi_i, \tau) \rho(|\pi_i|, \tau) \rho(\pi_i, \tau) \rho(|\pi_i|, \tau) \rho(\pi_i, \tau) \rho(|\pi_i|, \tau) \rho(\pi_i, \tau) \rho(|\pi_i|, \tau) \) for all \( \pi_{0:2t} \in \mathbb{I}_t \) and \( \tau \geq 0 \). Since \( \bar{\rho} \) is a \( K \) function, the FIE is RGAS by definition. With Lemma 1, this implies that the corresponding MHE under Assumptions 1 and 2 with a long enough horizon will be pRGAS.

If, in addition, \( \alpha(\cdot, 0) \) and \( \rho(\cdot, 0) \) are globally Lipschitz continuous at the origin, so will be \( \beta(\cdot, 0) \) defined above. Consequently, by Lemma 1, the FIE, which is RGAS, implies that the corresponding MHE with a sufficiently long horizon will be RGES. This completes the proof.

Remark 3: The three comments are given as follows.

1) The second conclusion of Theorem 1 claims RGES of a corresponding MHE but the FIE, as the proof is unclear for the latter, although FIE can be viewed as a limit of the MHE with \( T \to \infty \).
2) If the two \( K \) functions \( \alpha \) and \( \rho \) are both in exponential forms, then it follows from the proof of Theorem 1 that the FIE will also be RGES in addition to the corresponding MHE.
3) It can be shown that a general conclusion of RGAS (instead of RGES) of the MHE will be achieved if a looser Lipschitz continuity condition is satisfied. However, the analysis is rather involved and subject to the variant assumption of having different bound functions (as counterparts of \( \alpha, \rho, \beta \) adhering to different uncertainties indicated by \( x, w, v \). Due to space limit, the result is not presented here.

Next, we present a lemma indicating that Assumptions 3 and 4 do not impose special difficulty as there always exists a cost function satisfying both of them if the system is i-IOSS.

Lemma 3 (Satisfaction of Assumptions 3 and 4): If the system is i-IOSS as per (4), then, given any \( t \geq 0 \) and \( \rho \in \mathbb{K} \) satisfying \( \rho(\pi_i, \tau) \geq \alpha(2|\pi_i|, \tau) \) for all \( \pi_{0:2t} \in \mathbb{I}_t \) and \( \tau \geq 0 \), it is feasible to specify the cost function of FIE as

\[
\hat{V}_i(\pi_{0:2t}) := \max_{i \in \mathbb{I}_{0:2t}} \rho(|\pi_i|, t - \iota(\hat{\pi}_i) - 1)
\]

such that Assumptions 3 and 4 hold true. If, furthermore, the system is exp-i-IOSS, then the \( KL \) function \( \rho \) adopts a form as \( \rho(s, \tau) = c|\pi_i|b^{\tau} \) with certain \( c > 0 \) and \( b \in (0, 1) \) for all \( \tau \geq 0 \) and \( \hat{\pi}_i \) in the applicable domain.

Proof: Given the FIE cost function specified as per (16), Assumption 3 is automatically met. With \( \rho(\pi_i, \tau) \geq \alpha(2|\pi_i|, \tau) \) for all \( \pi_{0:2t} \in \mathbb{I}_t \) and \( \tau \geq 0 \), Assumption 4 is also met by Lemma 2. Therefore, the existence of such \( \rho \) is guaranteed whenever it is always feasible to let \( \rho(\pi_i, \tau) := \alpha(2|\pi_i|, \tau) \).

If the system is exp-i-IOSS with \( \alpha(\pi_i, \tau) := c'|\pi_i|b^{\tau} \) for certain \( c' > 0 \) and \( b \in (0, 1) \), then it is valid to let \( \rho(\pi_i, \tau) := \alpha(2|\pi_i|, \tau) \).

This completes the proof.

Lemma 3 indicates that a valid cost function can always be designed from the i-IOSS bound function \( \alpha \), and hence implies an important result in the following.

Corollary 1: There exists a cost function for FIE to be RGAS (or RGES) if and only if the system is i-IOSS (or exp-i-IOSS).

Proof: Sufficiency: By Lemma 3, the system being i-IOSS implies that the FIE admits a cost function such that Assumptions 3 and 4 hold true, which consequently implies that the FIE is RGAS by Theorem 1. When the system is exp-i-IOSS, the conclusion trivially strengthens to that the FIE is RGES by following the same approach of reasoning.

Necessity: The necessity in the RGAS case has been proved in existing literature, e.g., [1, Prop. 4.6] and [24, Prop. 2.4], while the proof in the RGES case is implied by the same proof there.

With Lemma 1, it follows immediately from Corollary 1 that there exists a cost function for MHE to be pRGAS if the system is i-IOSS and the associated uncertainties are bounded, and to be RGES if the system is exp-i-IOSS. The MHE will be RGAS if the i-IOSS system is such that the corresponding FIE satisfies a Lipschitz condition indicated in Lemma 1.

Remark 4: The conclusion of Corollary 1 coincides with a key finding reported in a latest paper [25], which had been submitted for review. The derivation approaches are, however, quite different. Here, we apply the reasoning approach of [2], focusing on developing more general conditions for robust stability of FIE, and the aforementioned conclusion appears as a corollary for an endeavor to understand the developed conditions. In contrast, Kneuer and Mueller [25] reached the conclusion by starting with a particular cost function, which is constructed directly from the i-IOSS property of the system and is not necessarily the only form admitted by our derived conditions. On the other hand, Kneuer and Mueller [25] also presented an admissible sum-based cost function, which is not covered in this work.

VI. CONCLUSION

This work proved a generic property that RGAS of FIE implies pRGAS of the corresponding MHE, which implements a sufficiently long horizon, and that the pRGAS strengthens to RGAS of the MHE if the FIE admits a certain global Lipschitz continuity property. With the revealed implication, sufficient conditions for the MHE to be pRGAS or RGAS were derived by first developing those for ensuring robust stability of the corresponding FIE. A particular realization of these conditions indicates that the system being i-IOSS is not only necessary but also sufficient to ensure the existence of an RGAS FIE. With the revealed implication, the sufficiency is also applicable to the existence of a pRGAS MHE if the associated uncertainties are bounded, and to the existence of an RGAS MHE if, in addition, a global Lipschitz continuity condition is satisfied. The established generic stability link paves the way for future stability analysis of MHE via deeper analysis of its corresponding FIE.
REFERENCES

[1] J. B. Rawlings, D. Q. Mayne, and M. Diehl, *Model Predictive Control: Theory, Computation, and Design*, 2nd ed. Santa Barbara, CA, USA: Nob Hill Publishing, 2020.

[2] W. Hu, “Robust stability of optimization-based state estimation,” 2017, arXiv:1702.01903, 2017.

[3] H. Michalska and D. Mayne, “Moving horizon observers,” in *Proc. Non-linear Control Syst. Des.*, 1993, pp. 185–190.

[4] K. R. Muske, J. B. Rawlings, and J. H. Lee, “Receding horizon recursive state estimation,” in *Proc. Amer. Control Conf.*., 1993, pp. 900–904.

[5] H. Michalska and D. Q. Mayne, “Moving horizon observers and observer-based control,” *IEEE Trans. Autom. Control*, vol. 40, no. 6, pp. 995–1006, Jun. 1995.

[6] C. Rao, J. Rawlings, and D. Mayne, “Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations,” *IEEE Trans. Autom. Control*, vol. 48, no. 2, pp. 246–258, Feb. 2003.

[7] A. Alessandri, M. Baglietto, and G. Battistelli, “Moving-horizon state estimation for nonlinear discrete-time systems: New stability results and approximation schemes,” *Automatica*, vol. 44, no. 7, pp. 1753–65, 2008.

[8] J. B. Rawlings and D. Q. Mayne, *Model Predictive Control: Theory and Design*. Madison, WI, USA: Nob Hill Publishing, 2009.

[9] A. Alessandri, M. Baglietto, G. Battistelli, and V. Zavala, “Advances in moving horizon estimation for nonlinear systems,” in *Proc. IEEE 49th Conf. Decis. Control*, 2010, pp. 5681–5688.

[10] J. Liu, “Moving horizon state estimation for nonlinear systems with bounded uncertainties,” *Chem. Eng. Sci.*, vol. 93, pp. 376–386, 2013.

[11] J. Rawlings and J. Li, “Optimization-based state estimation: Current status and some new results,” *J. Process Control*, vol. 22, pp. 1439–1444, 2012.

[12] E. Sontag and Y. Wang, “Output-to-state stability and detectability of nonlinear systems,” *Syst. Control Lett.*, vol. 29, no. 5, pp. 279–290, 1997.

[13] W. Hu, L. Xie, and K. You, “Optimization-based state estimation under bounded disturbances,” in *Proc. IEEE 54th Conf. Decis. Control*, 2015, pp. 6597–6602.

[14] L. Ji, J. B. Rawlings, W. Hu, A. Wynn, and M. Diehl, “Robust stability of moving horizon estimation under bounded disturbances,” *IEEE Trans. Autom. Control*, vol. 61, no. 11, pp. 3509–3514, Nov. 2016.

[15] M. A. Müller, “Nonlinear moving horizon estimation for systems with bounded disturbances,” in *Proc. Amer. Control Conf.*, 2016, pp. 883–888.

[16] M. A. Müller, “Nonlinear moving horizon estimation in the presence of bounded disturbances,” *Automatica*, vol. 79, pp. 306–314, 2017.

[17] D. A. Allan and J. B. Rawlings, “Moving horizon estimation,” in *Handbook of Model Predictive Control*. Cham, Switzerland: Birkhäuser, 2019, pp. 99–124.

[18] S. Knüfer and M. A. Müller, “Robust global exponential stability for moving horizon estimation,” in *Proc. IEEE 57th Conf. Decis. Control*, 2018, pp. 3477–3482.

[19] D. A. Allan, J. Rawlings, and A. R. Teel, “Nonlinear detectability and incremental input/output-to-state stability,” *SIAM J. Control Optim.*, vol. 59, no. 4, pp. 3017–3039, 2021.

[20] S. Knüfer and M. A. Müller, “Time-discounted incremental input/output-to-state stability,” in *Proc. IEEE 59th Conf. Decis. Control*, 2020, pp. 5394–5400.

[21] W. Hu, “Generic stability implication from full information estimation to moving-horizon estimation,” 2021, arXiv:2105.10125.

[22] L. Grüne and J. Pannek, “Nonlinear model predictive control: Theory and algorithms,” in *Proc. Nonlinear Model Predictive Control*, 2017, pp. 45–69.

[23] A. Aravkin, J. V. Burke, L. Ljung, A. Lozano, and G. Pillonetto, “Generalized Kalman smoothing: Modeling and algorithms,” 2016, arXiv:1609.06369.

[24] D. A. Allan, “A Lyapunov-like function for analysis of model predictive control and moving horizon estimation,” Ph.D. dissertation, Univ. Wisconsin–Madison, Madison, WI, USA, 2020.

[25] S. Kneuer and M. A. Mueller, “Nonlinear full information and moving horizon estimation: Robust global asymptotic stability,” 2021, arXiv:2105.02764.