TWCCC ⋆ Texas – Wisconsin – California Control Consortium

Technical report number 2020-01

Nonlinear Detectability and Incremental Input/output-to-state Stability

Douglas A. Allan, James B. Rawlings, Andrew R. Teel

July 4, 2020

Abstract

Incremental input/output-to-state stability (i-IOSS) is a popular characterization of detectability for nonlinear systems. For instance, it is known that any system that admits a robustly stable full-order observer (i.e., a system evolving in the same state space as the original system stabilized by output injection) must be i-IOSS. Nevertheless, there are many types of state estimator (such as moving horizon estimation and extended Kalman filtering) that are not such observers. Additionally, while other ISS-like properties have been characterized by storage functions with associated converse theorems, no converse theorem exists for i-IOSS. Here, we demonstrate that any system that admits a robustly stable state estimator must be i-IOSS, and, building on techniques developed for incremental input-to-state stability, provide an i-IOSS converse theorem. We also provide a result about changing supply rates for a general storage function, and apply it to an i-IOSS Lyapunov function as a corollary.

1 Introduction

One of the largest contributions Kalman and Bucy made in their seminal papers on linear filtering [18, 19] is not the design of a statistically optimal filter (though they did accomplish that), but the introductions of the concepts of controllability and observability from control and systems theory to the filtering literature. Kailath [17, p 152], reflecting on linear filtering, stated:

> As Kalman has often stressed [68] the major contribution of his work is not perhaps the actual filter algorithm, elegant and useful as it no doubt is, but

> the proof that under certain technical conditions called "controllability" and "observability," the optimum filter is "stable" or "robust" in the sense that the effects of initial errors and round-off and other computational errors will die out asymptotically. However, the known proofs of this result are somewhat difficult, and it is significant that only a small fraction of the vast literature on the Kalman filter deals with this problem.

However, it became clear that while controllability and observability are sufficient conditions for filter stability, they are not necessary conditions. More general conditions called detectability and stabilizability were quickly introduced for time invariant systems by Wonham [44], and were generalized to time varying systems by Anderson and Moore [5].

Shortly thereafter, observability was generalized to certain continuous-time bilinear systems by Brockett [9] and analytic systems by Sussmann [41]. Hermann and Krener [14] proposed *four* notions of observability, depending on whether all states can be distinguished from one another or just nearby states, and whether the system must make a large sojourn to distinguish states or whether they can be distinguished locally. They called systems for which local states can be distinguished by local measurements "locally weakly observable", and, for analytic systems, provided a sufficient condition for this property by a rank condition for the matrix of Lie derivatives, analogous to the well-known condition on the rank of the observability matrix for linear systems. They also provided a converse theorem, showing that analytic locally weakly observable systems satisfy that rank condition almost everywhere in their state spaces.

This program was extended to discrete-time polynomial systems by Sontag [35]. Results for time-sampled continuous-time systems were obtained by Aeyels [1] for smooth systems and Sontag [36] for analytic systems. Aeyels showed that almost all combinations of smooth state evolution equations with a finite number of fixed points, output equations, and sampling programs result in observable systems if $2n + 1$ samples are used, in which n is the dimension of the state space. It is somewhat reassuring that an arbitrary system is almost surely observable, but recall that almost all matrices are full rank and, if square, diagonalizable. Rank deficient matrices and defective matrices occur frequently in applications nevertheless. Sontag [36] showed that, for sampled analytic systems, n samples are sufficient. For discrete-time systems, Nijmeijer [27] proposed the concept of "strong local observability" for states that could be distinguished from their neighbors using n measurements and showed that this property is both implied by a matrix rank condition and implies that such a rank condition holds almost everywhere. One feature that many of these early notions of observability have in common is that, in order to distinguish two states, a certain input may be required to be applied to the system. In this case, we cannot expect to be able to design a controller and estimator independently of one another, as the separation principle permits us to do for linear systems.

The introduction of the notion of input-to-state stability (ISS) by Sontag [37] created a paradigm shift in how robust stability properties are considered. The notion of output-to-state stability (OSS) was proposed as a dual notion by Sontag and Wang [40] (though, as they note, inputs and outputs play different roles, and as a result proofs for OSS look very different than those for ISS). For an input-free linear system, it turns out that OSS is the same as detectability. However, for nonlinear systems, it is not sufficient to consider inputs and outputs independently, so they further propose input/output-to-state stability (IOSS). This property is sufficient for a nonlinear system to be "zero detectable" in the sense that if the inputs and outputs are bounded, then the state remains in a bounded neighborhood of the origin. However, this notion is *still* insufficient to serve as a nonlinear detectability assumption. Finally, Sontag and Wang [40] introduce incremental IOSS (i-IOSS), and they demonstrate that any system that admits a robust full-order state observer, i.e., a dynamical system evolving in the same state space as the system that converges to the system state under output injection, must also be i-IOSS. This notion of detectability was introduced to the optimization-based state estimation literature by Rawlings and Mayne [29, Ch. 4], and is now the standard in the field.

The characterization of ISS-like by storage functions and associated converse theorems, by which ISS-like systems are shown to admit such a storage function, has been a fruitful area of research. ISS converse theorems have appeared in Sontag and Wang [39] for continuous-time systems and Jiang and Wang [16] for discrete-time systems. A converse theorem for OSS appeared in Sontag and Wang [40], and one for IOSS appeared in Krichman et al. [24], both in continuous-time. Discrete-time converse theorems for OSS and IOSS both appeared in Cai and Teel [10]. All these results are formulated for continuous functions defined over $\mathbb{R}^n$ and $\mathbb{R}^m$, in which n is the dimension of the state space and m the dimension of the input space. An ISS converse theorem for discontinuous functions is given by Grüne and Kellett [13] and an IOSS converse theorem for closed, nonconvex domains is given in Allan and Rawlings [2].

Lyapunov and storage functions for incremental stability and ISS were proposed by Angeli [6, 7] in continuous time, however the converse theorem provided requires the inputs be limited to a compact set. The forward result that an incremental ISS Lyapunov function implies incremental ISS was extended to discrete time by Bayer et al. [8] and a converse theorem was given by Tran et al. [42], again, requiring inputs be limited to a compact set. Characterization of i-IOSS by an i-IOSS Lyapunov function was suggested by Müller [25] while a formal definition for an i-IOSS Lyapunov function was given in Allan and Rawlings [3]. The concepts of incremental passivity [28] and incremental dissipativity [33] are also related to i-IOSS.

Here, we first show that *any* system that admits a robustly stable estimator, which we define in terms of an input/output mapping, must be i-IOSS. We use the mapping definition of an estimator so that we can include optimization-based estimation techniques like full information estimation and moving horizon estimation. To apply this notion to systems that are being controlled by some external algorithm, we require the i-IOSS to be *uniform* (i-UIOSS) with respect to some time-varying parameter vector. Then, we show that *any* i-UIOSS system defined on a subset of a finite-dimensional vector space admits an i-IOSS Lyapunov function. It is possible, but cumbersome, to perform this proof from the traditional "asymptotic-gain" definition of i-UIOSS. However, we use a modified definition of i-UIOSS that reveals the stronger properties that are already inherent in i-UIOSS systems (at least those defined in subsets of finite-dimensional normed vector spaces). This "convolution-maximization" definition includes explicit discounting of past disturbances by a $\mathcal{KL}$ function. We include a proof that this form of i-UIOSS is equivalent to the asymptotic-gain definition in the appendix. Neither proof requires system regularity assumptions, but in the case that that the system evolution function and output function are $\mathcal{K}$ continuous (a variant of uniform continuity, equivalent for functions defined on convex subsets of normed vector spaces [11, p 232]), we show that the resulting i-UIOSS Lyapunov function is continuous. Finally, we include a result on changing supply rates for general storage functions, including i-UIOSS Lyapunov functions, that may be useful when applying them to model predictive control and moving horizon estimation (see, for example, Rawlings et al. [31, Sec. 2.4.4] for a similar application to an IOSS Lyapunov function for MPC).

Notation We denote the nonnegative integers by $\mathbb{I}_{\geq 0}$ and the nonnegative reals by $\mathbb{R}_{\geq 0}$. Integers ranging from a to b are denoted $\mathbb{I}_{a:b}$. For $x \in \mathbb{R}^n$, we denote its Euclidean norm by $|x|$. For a matrix $A \in \mathbb{R}^{n\times n}$, we denote its induced two norm as $|A|$. For some set $\mathcal{A}$, we denote point to set distance $|x|_{\mathcal{A}} := \inf_{y\in\mathcal{A}} |x-y|$. We denote a sequence of vectors $\{x(k)\}$ by $\mathbf{x}$. For a sequence $\mathbf{x}$ and singleton x,the operation $x \frown \mathbf{x}$ inserts x into the front of $\mathbf{x}$. The supremum norm of this sequence $\sup_{k\in\mathbb{I}_{\geq 0}} |x(k)| := \|\mathbf{x}\|$. We define $\|\mathbf{x}\|_{a:b} := \max_{k\in\mathbb{I}_{a:b}} |x(k)|$, and the finite subsequence from a to b as $\mathbf{x}_{a:b}$.

Recall that a function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class $\mathcal{K}$ if it is continuous, strictly increasing, and $\alpha(0) = 0$. It is of class $\mathcal{K}_\infty$ if, in addition, $\lim_{s\to\infty} \alpha(s) = \infty$. A function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{I}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class $\mathcal{KL}$ if $\beta(\cdot, k)$ is a $\mathcal{K}$ function for fixed $k \in \mathbb{I}_{\geq 0}$ and $\beta(s, \cdot)$ is a nonincreasing function that satisfies $\lim_{k\to\infty} \beta(s, k) = 0$ for fixed $s \in \mathbb{R}_{\geq 0}$.

Inspired by the max-plus algebra, for $a, b \in \mathbb{R}$ we define $a \oplus b := \max(a, b)$. In analogy with summation we write

$$\bigoplus_{j=0}^{k-1} a(j) := \max_{j\in\mathbb{I}_{0:k-1}} a(j).$$

Note that the operation $\oplus$ is associative, i.e., $(a \oplus b) \oplus c = a \oplus (b \oplus c)$, commutative, i.e., $a \oplus b = b \oplus a$, and satisfies a distributive property with strictly increasing functions. In particular, if $\sigma \in \mathcal{K}$, we have that $\sigma(a \oplus b) = \sigma(a) \oplus \sigma(b)$.

2 Robustly stable estimator

Consider a discrete-time system

$$x^+ = f(x, u, d) \qquad y = h(x), \tag{1}$$

in which $x \in \mathbb{X} \subseteq \mathbb{R}^n$ is the system state, $x^+ \in \mathbb{X}$ is the successor state, $u \in \mathbb{U} \subseteq \mathbb{R}^m$ is the system input, $d \in \mathbb{D} \subseteq \mathbb{R}^d$ is some time-varying parameter (or disturbance) vector, and $y \in \mathbb{Y} \subseteq \mathbb{R}^p$ is the system output. Note that we do *not* assume that the origin is a steady state, but merely assume that $\mathbb{X}$, $\mathbb{U}$, and $\mathbb{D}$ are nonempty. We denote the trajectory starting from x evolving with input sequence $\mathbf{u}$ and parameter sequence $\mathbf{d}$ as $\mathbf{x}(x, \mathbf{u}, \mathbf{d})$ and k^{th} element of this trajectory as $\phi(k; x, \mathbf{u}, \mathbf{d})$. We denote the k^{th} output of this trajectory by $y(k; x, \mathbf{u}, \mathbf{d})$ and denote the entire output sequence by $\mathbf{y}(x, \mathbf{u}, \mathbf{d})$. Where it is unambiguous, we usually abbreviate $\phi(k; x, \mathbf{u}, \mathbf{d})$ by $x(k)$ and $y(k; x, \mathbf{u}, \mathbf{d})$ by

$y(k)$. Because we are working with incremental stability properties, it is useful to use the abbreviated notation

$$\begin{aligned}
\Delta x(j) &:= \phi(j; x_1, \mathbf{u}_1, \mathbf{d}) - \phi(j; x_2, \mathbf{u}_2, \mathbf{d}) & \Delta \mathbf{x} &:= \mathbf{x}(x_1, \mathbf{u}_1, \mathbf{d}) - \mathbf{x}(x_2, \mathbf{u}_2, \mathbf{d}) \\
\Delta u(j) &:= u_1(j) - u_2(j) & \Delta \mathbf{u} &:= \mathbf{u}_1 - \mathbf{u}_2 \\
\Delta y(j) &:= y(j; x_1, \mathbf{u}_1, \mathbf{d}) - y(j; x_2, \mathbf{u}_2, \mathbf{d}) & \Delta \mathbf{y} &:= \mathbf{y}(x_1, \mathbf{u}_1, \mathbf{d}) - \mathbf{y}(x_2, \mathbf{u}_2, \mathbf{d})
\end{aligned}$$

in which the dependence of these increments on x_1, x_2, $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{d}$ is suppressed for brevity.

We are interested in i-UIOSS as a definition of nonlinear detectability. For a detectable linear system

$$x^+ = Ax + Bu \qquad y = Cx \tag{2}$$

it is known that there exists some matrix L such that $A - LC$ is Schur stable, i.e., has all its eigenvalues strictly within the unit disk. We can thus use L to construct an observer such that

$$\hat{x}^+ = A\hat{x} + L(y^m - C\hat{x}) + Bu^f,$$

in which u^f is a forecast of the input u which is off by some forecasting error $w := u - u_f$ and y^m is the measurement y after being corrupted by some output noise $v := y^m - y$. The estimate error $e(k) := x(k) - \hat{x}(k)$ then evolves as

$$e^+ = (A - LC)e - Lv + Bw.$$

We can express the evolution of this linear system as a convolution sum

$$e(k) = (A - LC)^k e(0) - \sum_{j=0}^{k-1} (A - LC)^{k-j-1} Lv(j) + \sum_{j=0}^{k-1} (A - LC)^{k-j-1} Bw(j)$$

that, after using the well-known fact that any Schur stable matrix admits an upper bound $\left|(A - LC)^k\right| \leq K\eta^k$ for some $K > 0$ and $\eta \in (0, 1)$, implies we have the upper bound

$$|e(k)| \leq K\eta^k |e(0)| + \sum_{j=0}^{k-1} K |L| \eta^{k-j-1} |v(j)| + \sum_{j=0}^{k-1} K |B| \eta^{k-j-1} |w(j)| \tag{3}$$

for all initial conditions and disturbance sequences.

Sontag and Wang [40] considered a direct nonlinear equivalent to such an observer. They defined a (robustly stable) full-order state observer as a dynamical system

$$\hat{x}^+ = g(\hat{x}, u^f, y^m) \tag{4}$$

evolving in the same state-space $\mathbb{X}$ as the true system state, that admits an upper bound

$$|e(k)| \leq \beta(|e(0), k|) \oplus \gamma_u(\|\mathbf{w}\|_{0:k-1}) \oplus \gamma_v(\|\mathbf{v}\|_{0:k-1}) \tag{5}$$

for some $\beta \in \mathcal{KL}$ and $\gamma_u, \gamma_y \in \mathcal{K}$. However, though designing such observers is an active area of research, there are many methods of nonlinear state estimation that do not fit into such a framework. For example, the extended Kalman filter (EKF) has a state that consists of not only an estimate of the system state, but also an estimate of the covariance of the system state. As a result, it evolves in a state space of higher dimension than $\mathbb{X}$. Other optimization-based state estimators such as full information estimation (FIE) and moving horizon estimation (MHE) have no convenient state-space representation. As a result, we follow the suggestion of Sontag and Wang [40, Remark 25] by defining state estimators in terms of input/output maps.

Definition 1 (State Estimator)**.** A *state estimator* is a sequence of functions $\Psi_k : \mathbb{X} \times \mathbb{U}^k \times \mathbb{D}^k \times \mathbb{Y}^k \to \mathbb{X}$ defined for all $k \in \mathbb{I}_{\geq 0}$. We define state estimates

$$\hat{x}(k) = \Psi_k(\overline{x}, \mathbf{u}^f_{0:k-1}, \mathbf{d}_{0:k-1}, \mathbf{y}^m_{0:k-1})$$

in which $\mathbf{u}^f \in \mathbb{U}^\infty$ is a forecast of inputs, $\mathbf{y}^m$ is the measured output sequence, and $\mathbf{d}$ is the true time-varying parameter sequence entering into the system. ∎

Equation (5) does not adequately characterize robust stability for such a general class of estimators: there is no longer the implication that state estimates converge if the disturbance sequences $\mathbf{w}$ and $\mathbf{v}$ converge to zero. When a full-order state estimator is used, the bound eq. (5) can be repeatedly applied to move the effects of disturbances from the asymptotic gains, which remain constant, to the state, which decays because of the $\mathcal{KL}$ function. If a more general estimator in the class defined by theorem 1 is used, however, the initial state has a unique role, because

$$\begin{aligned}\Psi_{2k}(\overline{x}, \mathbf{u}^f_{0:2k-1}, \mathbf{d}_{0:2k-1}, \mathbf{y}^m_{0:2k-1}) \neq \Psi_k\Big(&\Psi_k(\overline{x}, \mathbf{u}^f_{0:k-1}, \mathbf{d}_{0:k-1}, \mathbf{y}^m_{0:k-1}),\\ &\mathbf{u}^f_{k:2k-1}, \mathbf{d}_{k:2k-1}, \mathbf{y}^m_{k:2k-1}\Big)\end{aligned}$$

in general (as is the case for FIE). We therefore modify eq. (5), making it more closely resemble eq. (3).

Definition 2 (Robust stability)**.** Suppose that $\mathbf{w} \coloneqq \mathbf{u} - \mathbf{u}^f$ and that $\mathbf{v} \coloneqq \mathbf{y}^m - \mathbf{y}$, in which $\mathbf{u} \in \mathbb{U}^\infty$ is the true input sequence entering into the system and $\mathbf{y}$ is the true sequence of outputs without noise. A state estimator (Ψ_k) is *robustly stable* if there exist $\rho_x, \rho_w, \rho_v \in \mathcal{KL}$ such that

$$\begin{aligned}|x(k) - \hat{x}(k)| \leq \rho_x(|x(0) - \overline{x}|, k) &\oplus \bigoplus_{j=0}^{k-1} \rho_w(|w(j)|, k-j-1)\\ &\oplus \bigoplus_{j=0}^{k-1} \rho_v(|v(j)|, k-j-1)\end{aligned}$$

for all $k \in \mathbb{I}_{\geq 0}$, $\mathbf{u}, \mathbf{u}^f \in \mathbb{U}^\infty$, $\mathbf{d} \in \mathbb{D}^\infty$, and $\mathbf{y}, \mathbf{y}^m$. ∎

In contrast to the "asymptotic-gain" formulation used in eq. (5), we instead explicitly discount past disturbances in this "convolution-maximization" formulation of robust stability. In comparison to eq. (3), we are using general $\mathcal{KL}$ functions instead of one of the specific form $\beta(s,k) \coloneqq Ks\eta^k$ for $K > 0$ and $\eta \in (0,1)$, and maximizing over the interval $\mathbb{I}_{0:k-1}$ rather than summing. We would expect the former difference when passing from the linear to nonlinear cases, and the latter difference occurs because there is no guarantee an arbitrary $\mathcal{KL}$ function sufficiently discounts a bounded sequence such that it is summable. For an exponential function, we have that

$$\bigoplus_{j=0}^{k-1} z(j)\eta^{k-j-1} \leq \sum_{j=0}^{k-1} z(j)\eta^{k-j-1} \leq \frac{1}{1-\eta^{1/2}} \bigoplus_{j=0}^{k-1} z(j)\eta^{\frac{k-j-1}{2}}$$

so the two operations are in a sense equivalent (compare, for example, the definition of *exponential* i-IOSS used by Knüfer and Müller [23] using convolution *sums*). We next define i-UIOSS along similar lines.

Definition 3 (Incremental uniform input/output-to-state stability)**.** A system eq. (1) is i-UIOSS if there exist $\beta_x, \beta_u, \beta_y \in \mathcal{KL}$ such that

$$\begin{aligned} |\Delta x(k)| \leq \beta_x(|\Delta x(0)|, k) &\oplus \bigoplus_{j=0}^{k-1} \beta_u(|\Delta u(j)|, k-j-1) \\ &\oplus \bigoplus_{j=0}^{k-1} \beta_y(|\Delta y(j)|, k-j-1), \end{aligned} \tag{6}$$

for all $x_1, x_2 \in \mathbb{X}$, all $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{U}^\infty$, and $\mathbf{d} \in \mathbb{D}^\infty$. ■

Traditionally, i-UIOSS has been formulated in terms of a single $\mathcal{KL}$ function $\beta(\cdot)$ and $\mathcal{K}$ function asymptotic gains $\gamma_u(\cdot), \gamma_y(\cdot)$ such that

$$|\Delta x(k)| \leq \beta(|\Delta x(0)|, k) \oplus \gamma_u(\|\Delta \mathbf{u}\|_{0:k-1}) \oplus \gamma_y(\|\Delta \mathbf{y}\|_{0:k-1}), \tag{7}$$

but we have found it more insightful to formulate these properties in a form where there is explicit discounting of input and output differences. It is straightforward to find a bound like eq. (7) from one like eq. (6), but we know of no easy method to derive a bound like eq. (6) from eq. (7). Nevertheless, they are equivalent.

Proposition 4. *A system eq.* (1) *admits an upper bound of the form eq.* (7) *for all $x_1, x_2 \in \mathbb{X}$, all $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{U}^\infty$, and $\mathbf{d} \in \mathbb{D}^\infty$ if and only if it is i-UIOSS as characterized in theorem 3.*

We defer proof of this proposition to the appendix, because it requires pushing the techniques developed by Krichman et al. [24], Cai and Teel [10], and Allan and Rawlings [3] to their limit. However, study of i-UIOSS as characterized in theorem 3 can be motivated independently of its equivalence to i-UIOSS as previously defined in the literature. We next show that any system that admits a robustly stable state estimator must satisfy theorem 3. The key insight required is that for any combination of initial guess, input sequence, parameter sequence, and output sequence that is feasible, i.e., obeys the state evolution equation eq. (1), the robust estimator must return the corresponding state sequences.

Proposition 5. *A system eq.* (1) *admits a robustly stable estimator* (Ψ_k) *only if it is i-UIOSS.*

Proof. Let x_1 and x_2 be two arbitrary initial states, with corresponding input sequences $\mathbf{u}_1$ and $\mathbf{u}_2$ and one single time-varying parameter sequence $\mathbf{d}$ used for both trajectories. These sequences give rise to state sequences $\mathbf{x}_1$ and $\mathbf{x}_2$ and output sequences $\mathbf{y}_1$ and $\mathbf{y}_2$. Suppose we use $\overline{x} = x_i$, $\mathbf{u}^f = \mathbf{u}_i$, and $\mathbf{y}^m = \mathbf{y}_i$ to estimate $x_i(k)$ (and thus have $\mathbf{w} = \mathbf{0}$ and $\mathbf{v} = \mathbf{0}$). Then we can apply theorem 2 to obtain:

$$\begin{aligned} |x_i(k) - \hat{x}(k)| \leq \rho_x(|x_i(0) - x_i(0)|\,, k) &\oplus \bigoplus_{j=0}^{k-1} \rho_w(|0|\,, k-j-1) \\ &\oplus \bigoplus_{j=0}^{k-1} \rho_v(|0|\,, k-j-1) \end{aligned}$$

for $i = 1, 2$. Thus, we have that

$$\Psi_k(x_1(0), \mathbf{u}_1, \mathbf{d}, \mathbf{y}_1) = x_1(k) \qquad \Psi_k(x_2(0), \mathbf{u}_2, \mathbf{d}, \mathbf{y}_2) = x_2(k) \tag{8}$$

for all $k \in \mathbb{I}_{\geq 0}$. Next, consider what would happen if we used $\overline{x} = x_2$, $\mathbf{u}^f = \mathbf{u}_2$, and $\mathbf{y}^m = \mathbf{y}_2$ in order to estimate $x_1(k)$. In this case, we have that $\mathbf{w} = \mathbf{u}_1 - \mathbf{u}_2$ and $\mathbf{v} = \mathbf{y}_2 - \mathbf{y}_1$. We can apply theorem 2 to obtain:

$$|\hat{x}(k) - x_1(k)| \leq \rho_x(|\Delta x(0)|\,, k) \oplus \bigoplus_{j=0}^{k-1} \rho_u(|\Delta u(j)|\,, k-j-1) \oplus \bigoplus_{j=0}^{k-1} \rho_y(|\Delta y(j)|\,, k-j-1).$$

Equation (8) shows that $\hat{x}(k) = x_2(k)$ in this case. Therefore, because $x_1, x_2, \mathbf{u}_1, \mathbf{u}_2$ and $\mathbf{d}$ were arbitrary, the system eq. (1) is i-UIOSS. □

Therefore i-UIOSS is a *necessary* condition for a system to admit a robustly stable estimator. It is unknown whether it is also a *sufficient* condition, as detectability is for linear systems.

We seek to characterize i-UIOSS through a storage function and dissipation inequality.

Definition 6 (i-UIOSS Lyapunov function)**.** A function $V : \mathbb{X} \times \mathbb{X} \to \mathbb{R}_{\geq 0}$ is called an (exponential-decrease) i-IOSS Lyapunov function if there exist $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$, $\sigma_u, \sigma_y \in \mathcal{K}$, and $\eta \in (0, 1)$ such that

$$\alpha_1(|x_1 - x_2|) \leq V(x_1, x_2) \leq \alpha_2(|x_1 - x_2|) \tag{9}$$

$$\begin{aligned} V(f(x_1, u_1, d), f(x_2, u_2, d)) \leq{}& \eta V(x_1, x_2) + \sigma_u(|u_1 - u_2|) \\ &+ \sigma_y(|h(x_1) - h(x_2)|) \end{aligned} \tag{10}$$

for all $x_1, x_2 \in \mathbb{X}$, $u_1, u_2 \in \mathbb{U}$, and $d \in \mathbb{D}$. ■

A dissipation inequality of the form

$$\begin{aligned}V(f(x_1,u_1,d),f(x_2,u_2,d)) \leq& V(x_1,x_2) - \alpha_3(|x_1-x_2|) + \sigma_u(|u_1-u_2|)\\ &+ \sigma_y(|h(x_1)-h(x_2)|)\end{aligned}$$

for some $\mathcal{K}_\infty$ function $\alpha_3(\cdot)$ also works to define an i-UIOSS Lyapunov function, and can be easily derived from an exponential-decrease Lyapunov function.

A proof that an (non-uniform) i-IOSS Lyapunov function implies that a system is i-IOSS has appeared previously in [3, Proposition 5, Remark 6], and the addition of the time-varying parameter vector d to the problem does not impact the proof in a meaningful way. The converse implication is the focus of this paper.

3 i-UIOSS Converse Theorem

The key tool for this converse theorem, as well as most other works of this type, is Sontag's $\mathcal{KL}$ function lemma.

Proposition 7 (Sontag [38, Proposition 7])**.** *For every $\beta \in \mathcal{KL}$ there exist $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$ such that*

$$\alpha_1(\beta(s,k)) \leq \alpha_2(s)e^{-k}$$

for all $s \in \mathbb{R}_{\geq 0}$ and $k \in \mathbb{I}_{\geq 0}$.

Previous works for (non-incremental) UIOSS, like Cai and Teel [10] and Allan and Rawlings [2], proceed by defining an asymptotically stable difference inclusion, obtaining a Lyapunov function for *that* system, then showing that it also serves as an UIOSS Lyapunov for the original system. This procedure can be modified to work for the incremental case (and is, in fact, how we demonstrate that asymptotic-gain i-UIOSS is equivalent to convolution-maximization i-UIOSS in the appendix), but the resulting difference inclusion is not necessarily compact-valued, even for most linear systems. That precludes application of the powerful results regarding converse theorems for difference inclusions in Kellett and Teel [20, 21, 22], which the results in Cai and Teel [10] and Allan and Rawlings [2] depended upon.

The new approach we use finds its origins in Angeli [6, 7], as well as application of Sontag's lemma as in, for example, Kellett and Teel [21]. The major innovation in this approach is leveraging the convolution-maximization form of i-UIOSS to deal with inputs and outputs. We proceed in two stages. First we generate an i-UIOSS Lyapunov function without any regularity assumptions for the system eq. (1) beyond i-UIOSS itself. Then, upon assuming the system is $\mathcal{K}$ continuous, we show that the resulting i-UIOSS Lyapunov function is continuous.

Theorem 8. *A system eq.* (1) *is i-UIOSS if and only if it admits an (exponential-decrease) i-UIOSS Lyapunov function.*

Proof. The proof that an i-IOSS Lyapunov function implies i-IOSS provided in Allan and Rawlings [3, Proposition 5, Remark 6] requires negligible modification to account for the time-varying parameter vector d. The converse implication is much more involved.

We start by applying theorem 7 to $\beta_x(\cdot)$, $\beta_u(\cdot)$, and $\beta_y(\cdot)$ to find $\alpha_{1,x}, \alpha_{2,x}, \alpha_{1,u}, \alpha_{2,u}, \alpha_{1,y}, \alpha_{2,y} \in \mathcal{K}_\infty$ such that

$$\begin{aligned}
\alpha_{1,x}(\beta_x(s,k)) &\leq \alpha_{2,x}(s)e^{-k} \\
\alpha_{1,u}(\beta_u(s,k)) &\leq \alpha_{2,u}(s)e^{-k} \\
\alpha_{1,y}(\beta_y(s,k)) &\leq \alpha_{2,y}(s)e^{-k}.
\end{aligned}$$

Let $\lambda := e^{-1}$ and $\alpha(s) := \min(\alpha_{1,x}(s), \alpha_{1,u}(s), \alpha_{1,y}(s))$. We then have that

$$\begin{aligned}
\alpha(\beta_x(s,k)) &\leq \alpha_x(s)\lambda^k \\
\alpha(\beta_u(s,k)) &\leq \alpha_u(s)\lambda^k \\
\alpha(\beta_y(s,k)) &\leq \alpha_y(s)\lambda^k,
\end{aligned}$$

in which we have suppressed the subscript 2 on $\alpha_x(\cdot)$, $\alpha_u(\cdot)$, and $\alpha_y(\cdot)$ for brevity. We now define the i-UIOSS Lyapunov function candidate

$$\begin{aligned}
V(x_1, x_2) := \sup_{k \in \mathbb{I}_{\geq 0}, \mathbf{u}_1, \mathbf{u}_2, \mathbf{d}} \lambda^{-k/2} \Bigg[\alpha(|\Delta x(k)|) &- \sum_{j=0}^{k-1} 2\alpha_u(|\Delta u(j)|)\lambda^{k-j-1} \\
&- \sum_{j=0}^{k-1} \alpha_y(|\Delta y(j)|)\lambda^{k-j-1} \Bigg], \qquad (11)
\end{aligned}$$

in which the extra factor of two in the input term is intended, and is used in the proof for continuity. Note that we have

$$\begin{aligned}
\alpha(|\Delta x(k)|) \leq &\alpha\big(\beta_x(|\Delta x(0)|, k)\big) \oplus \alpha\left(\bigoplus_{j=0}^{k-1} \beta_u(|\Delta u(j)|, k-j-1) \right) \\
&\oplus \alpha\left(\bigoplus_{j=0}^{k-1} \beta_y(|\Delta y(j)|, k-j-1) \right) \\
\leq &\alpha_x(|\Delta x(0)|)\lambda^k \oplus \bigoplus_{j=0}^{k-1} \alpha_u(|\Delta u(j)|)\lambda^{k-j-1} \oplus \bigoplus_{j=0}^{k-1} \alpha_y(|\Delta y(j)|)\lambda^{k-j-1} \\
\leq &\alpha_x(|\Delta x(0)|)\lambda^k + \sum_{j=0}^{k-1} \alpha_u(|\Delta u(j)|)\lambda^{k-j-1} + \sum_{j=0}^{k-1} \alpha_y(|\Delta y(j)|)\lambda^{k-j-1}, \qquad (12)
\end{aligned}$$

and, as a result, we have that

$$\begin{aligned}
V(x_1, x_2) &\leq \sup_{k \in \mathbb{I}_{\geq 0}, \mathbf{u}_1, \mathbf{u}_2, \mathbf{d}} \lambda^{-k/2} \left(\alpha_x(|\Delta x(0)|)\lambda^k - \sum_{j=0}^{k-1} \alpha_u(|\Delta u(j)|)\lambda^{k-j-1} \right) \\
&\leq \sup_{k \in \mathbb{I}_{\geq 0}, \mathbf{u}_1, \mathbf{u}_2, \mathbf{d}} \alpha_x(|\Delta x(0)|)\lambda^{k/2} \\
&= \alpha_x(|x_1(0) - x_2(0)|), \qquad (13)
\end{aligned}$$

which provides the $\mathcal{K}_\infty$ upper bound on $V(x_1, x_2)$ indicated in eq. (9) with $\alpha_2(\cdot) \coloneqq \alpha_x(\cdot)$. We can establish the corresponding $\mathcal{K}_\infty$ lower bound with $\alpha_1(\cdot) \coloneqq \alpha(\cdot)$ by noting

$$\begin{aligned}\sup_{k\in\mathbb{I}_{\geq 0},\mathbf{u}_1,\mathbf{u}_2,\mathbf{d}} \lambda^{-k/2}&\left[\alpha(|\Delta x(k)|) - \sum_{j=0}^{k-1} 2\alpha_u(|\Delta u(j)|)\lambda^{k-j-1} - \sum_{j=0}^{k-1} \alpha_y(|\Delta y(j)|)\lambda^{k-j-1}\right] \\ &\geq \lambda^0\left[\alpha(|\Delta x(0)|) - \sum_{j=0}^{-1} 2\alpha_u(|\Delta u(j)|)\lambda^{j} - \sum_{j=0}^{-1} \alpha_y(|\Delta y(j)|)\lambda^{j}\right] \\ &= \alpha(|x_1(0) - x_2(0)|). \end{aligned} \tag{14}$$

We now derive the dissipation inequality eq. (10). For fixed $x_1, x_2 \in \mathbb{X}$, $u_1, u_2 \in \mathbb{U}$, and $d \in \mathbb{D}$, let $x_1^+ \coloneqq f(x_1, u_1, d)$ and $x_2^+ \coloneqq f(x_2, u_2, d)$. For every $\varepsilon > 0$, there exist $\mathbf{u}_1^*$, $\mathbf{u}_2^*$, $\mathbf{d}^*$, and $k^* \in \mathbb{I}_{\geq 0}$ such that

$$\begin{aligned} V(x_1^+, x_2^+) \leq \varepsilon + \lambda^{-k^*/2}\Bigg[&\alpha\left(\left|\phi(k^*; x_1^+, \mathbf{u}_1^*, \mathbf{d}^*) - \phi(k^*; x_2^+, \mathbf{u}_2^*, \mathbf{d}^*)\right|\right) \\ &- \sum_{j=0}^{k^*-1} 2\alpha_u(|u_1^*(j) - u_2^*(j)|)\lambda^{k^*-j-1} \\ &- \sum_{j=0}^{k^*-1} \alpha_y\left(\left|y(j; x_1^+, \mathbf{u}_1^*, \mathbf{d}^*) - y(j; x_2^+, \mathbf{u}_2^*, \mathbf{d}^*)\right|\right)\lambda^{k^*-j-1}\Bigg]. \end{aligned}$$

By considering trajectories beginning from x_1 and x_2, we have that

$$\begin{aligned} \phi(k^*; x_1^+, \mathbf{u}_1^*, \mathbf{d}^*) &= \phi(k^*+1; x_1, u_1 ^\frown \mathbf{u}_1^*, d ^\frown \mathbf{d}^*) \\ \phi(k^*; x_2^+, \mathbf{u}_2^*, \mathbf{d}^*) &= \phi(k^*+1; x_2, u_2 ^\frown \mathbf{u}_1^*, d ^\frown \mathbf{d}^*) \end{aligned}$$

in which the operation $s ^\frown \mathbf{s}$ inserts s at the beginning of the sequence $\mathbf{s}$. For brevity, let $\tilde{\mathbf{d}} \coloneqq d ^\frown \mathbf{d}^*$ and

$$\begin{aligned} &\tilde{\mathbf{u}}_1 \coloneqq u_1 ^\frown \mathbf{u}_1^* && \tilde{x}_1(j) \coloneqq \phi(j; x_1, \tilde{\mathbf{u}}_1, \tilde{\mathbf{d}}) && \tilde{y}_1(j) \coloneqq y(j; x_1, \tilde{\mathbf{u}}_1, \tilde{\mathbf{d}}) \\ &\tilde{\mathbf{u}}_2 \coloneqq u_2 ^\frown \mathbf{u}_2^* && \tilde{x}_2(j) \coloneqq \phi(j; x_2, \tilde{\mathbf{u}}_2, \tilde{\mathbf{d}}) && \tilde{y}_2(j) \coloneqq y(j; x_2, \tilde{\mathbf{u}}_2, \tilde{\mathbf{d}}). \end{aligned}$$

We thus have that

$$\begin{aligned}
V(x_1^+, x_2^+) \leq & \varepsilon + \lambda^{-k^*/2}\Bigg[\alpha(|\Delta\tilde{x}(k^*+1)|) \\
& - \sum_{j=1}^{k^*} 2\alpha_u(|\Delta\tilde{u}(j)|)\lambda^{k^*-j} - \sum_{j=1}^{k^*}\alpha_y(|\Delta\tilde{y}(j)|)\lambda^{k^*-j}\Bigg] \\
= & \varepsilon + \sqrt{\lambda}\Bigg(\lambda^{-(k^*+1)/2}\Bigg[\alpha(|\Delta\tilde{x}(k^*+1)|) - \sum_{j=1}^{k^*} 2\alpha_u(|\Delta\tilde{u}(j)|)\lambda^{k^*-j} \\
& - 2\lambda^{k^*}\alpha_u(|\Delta u(0)|) + 2\lambda^{k^*}\alpha_u(|\Delta u(0)|) \\
& - \sum_{j=1}^{k^*}\alpha_y(|\Delta\tilde{y}(j)|)\lambda^{k^*-j} - \lambda^{k^*}\alpha_y(|\Delta y(0)|) + \lambda^{k^*}\alpha_y(|\Delta y(0)|)\Bigg]\Bigg) \\
= & \varepsilon + \sqrt{\lambda}\Bigg(\lambda^{-(k^*+1)/2}\Bigg[\alpha(|\Delta\tilde{x}(k^*+1)|) \\
& - \sum_{j=0}^{k^*} 2\alpha_u(|\Delta\tilde{u}(j)|)\lambda^{k^*-j} - \sum_{j=0}^{k^*}\alpha_y(|\Delta\tilde{y}(j)|)\lambda^{k^*-j}\Bigg]\Bigg) \\
& + 2\lambda^{k^*/2}\alpha_u(|\Delta u(0)|) + \lambda^{k^*/2}\alpha_y(|\Delta y(0)|) \\
\leq & \varepsilon + \sqrt{\lambda}V(x_1, x_2) + 2\alpha_u(|\Delta u(0)|) + \alpha_y(|\Delta y(0)|),
\end{aligned}$$

in which the last step follows because $k^*+1, \tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2, \tilde{\mathbf{d}}$ is feasible for the optimization that produces $V(x_1, x_2)$. Note that, because we have removed all terms that depend on $\mathbf{u}_1^*$, $\mathbf{u}_2^*$, $\mathbf{d}^*$, and k^*, ε is arbitrary. We can thus take the limit as $\varepsilon \to 0$ to obtain

$$\begin{aligned}
V(x_1^+, x_2^+) &\leq \sqrt{\lambda}V(x_1, x_2) + 2\alpha_u(|u_1 - u_2|) + \alpha_y(|\tilde{y}_1(0) - \tilde{y}_2(0)|) \\
&= \sqrt{\lambda}V(x_1, x_2) + 2\alpha_u(|u_1 - u_2|) + \alpha_y(|h(x_1) - h(x_2)|).
\end{aligned}$$

This equation fulfills (10) with exponential decrease factor $\eta = \sqrt{\lambda}$ and $\mathcal{K}$ function supply rates $\sigma_u(\cdot) = 2\alpha_u(\cdot)$, and $\sigma_y(\cdot) = \alpha_y(\cdot)$. Thus $V(\cdot)$ is an (exponential-decrease) i-UIOSS Lyapunov function. □

Obtaining an i-UIOSS Lyapunov function by itself may be useful for applications, but continuity may be necessary for robustness results. First, we need a suitable continuity assumption for the underlying system eq. (1).

Assumption 9 ($\mathcal{K}$ Continuity)**.** Both $f(\cdot,\cdot,d)$ and $h(\cdot)$ are $\mathcal{K}$-continuous, the former uniformly in d, i.e., there exist $\sigma_f, \sigma_h \in \mathcal{K}$ such that

$$\begin{aligned}
|f(x_1, u_1, d) - f(x_2, u_2, d)| &\leq \sigma_f(|(x_1, u_1) - (x_2, u_2)|) \\
|h(x_1) - h(x_2)| &\leq \sigma_h(|x_1 - x_2|)
\end{aligned}$$

for all $x_1, x_2 \in \mathbb{X}$, all $u_1, u_2 \in \mathbb{U}$, and $d \in \mathbb{D}$.

As mentioned in the introduction, while $\mathcal{K}$ continuity and uniform continuity are, in general, distinct properties, they are identical in many cases of interest. Freeman and Kokotovic [11, p 232]) show that they coincide for functions defined on convex subsets of normed vector spaces, and [4, Proposition 20] implies that they coincide for compact subsets of $\mathbb{R}^n$ as well.

Remark 10. Because we define i-IOSS in predictor form (i.e., use outputs $(y(0), y(1), \ldots, y(k-1))$ but omit $y(k)$), we have the inequality

$$\left|x_1^+ - x_2^+\right| \leq \beta_x(|x_1 - x_2|, 1) + \beta_u(|u_1 - u_2|, 0) + \beta_y(|h(x_1) - h(x_2)|, 0).$$

If the function $h(\cdot)$ is $\mathcal{K}$-continuous, as occurs in the common case where a subset of states are measured, then theorem 9 is satisfied.

Theorem 9 can be used in order to show that $V(\cdot)$ as defined in theorem 8 is continuous.

Theorem 11. *The function $V(\cdot)$, as defined in eq. (11), is continuous if theorem 9 holds.*

The proof of this theorem, however, requires several minor results. The first is a tool used by Angeli [7], but never explicitly stated or named.

Proposition 12 (Quadrille inequality)**.** *For vectors $u_1, u_2, v_1, v_2 \in \mathbb{R}^n$, we have that*

$$\big|\, |u_1 - u_2| - |v_1 - v_2| \,\big| \leq |u_1 - v_1| + |u_2 - v_2|.$$

Proof. We have that

$$\big|\, |u_1 - u_2| - |v_1 - v_2| \,\big| \leq |u_1 - u_2 + v_2 - v_1|$$

by the reverse triangle inequality and

$$|u_1 - u_2 + v_2 - v_1| \leq |u_1 - v_1| + |u_2 - v_2|$$

by the regular triangle inequality. □

This next proposition generates upper bounds on $\|\Delta \mathbf{x}\|_{0:k}$ and $\|\Delta \mathbf{y}\|_{0:k}$ based on $|\Delta x|$ and $\|\Delta \mathbf{u}\|_{0:k-1}$. Notably, these bounds are uniform in x_1, x_2 and $\mathbf{u}_1, \mathbf{u}_2$ so long as $|\Delta x|$ and $\|\Delta \mathbf{u}\|_{0:k-1}$ are uniformly bounded above.

Proposition 13. *Let $\overline{s}_0 > 0$, $\overline{s}_u > 0$, and $\bar{k} \in \mathbb{I}_{\geq 0}$. If theorem 9 holds, then there exist some $\overline{s}_x(\overline{s}_0, \overline{s}_u, \bar{k}) \geq 0$ and $\overline{s}_y(\overline{s}_0, \overline{s}_u, \bar{k}) \geq 0$ such that*

$$|\Delta x(k)| \leq \overline{s}_x(\overline{s}_0, \overline{s}_u, \bar{k})$$
$$|\Delta y(k)| \leq \overline{s}_y(\overline{s}_0, \overline{s}_u, \bar{k})$$

for all $k \in \mathbb{I}_{0:\bar{k}}$ and any $x_1, x_2 \in \mathbb{X}$ such that $|\Delta x| \leq \overline{s}_0$ and any $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{U}^\infty$ such that $\|\Delta \mathbf{u}\| \leq \overline{s}_u$.

Proof. Let $\sigma_f, \sigma_h \in \mathcal{K}_\infty$ come from theorem 9. We have that

$$\begin{aligned} |f(x_1,u_1,d) - f(x_2,u_2,d)| &\le \sigma_f(|(\Delta x, \Delta u)|) \\ &= \sigma_f(|(\Delta x, 0) + (0, \Delta u)|) \\ &\le \sigma_f(|\Delta x| + |\Delta u|) \\ &\le \sigma_f(2\,|\Delta x|) \oplus \sigma_f(2\,|\Delta u|) \\ &\coloneqq \tilde{\sigma}_f(|\Delta x|) \oplus \tilde{\sigma}_f(|\Delta u|) \end{aligned}$$

for all $x_1, x_2 \in \mathbb{X}$ and $u_1, u_2 \in \mathbb{U}$. By recursively applying this equation, we obtain

$$|\Delta x(k)| \le \tilde{\sigma}_f^k(|\Delta x(0)|) \oplus \bigoplus_{j=0}^{k-1} \tilde{\sigma}_f^{k-j-1}(|\Delta u(j)|).$$

Without loss of generality, assume that $\tilde{\sigma}_f(s) \ge s$ for all $s \in \mathbb{R}_{\ge 0}$. We then have that $\tilde{\sigma}_f^k(s) \le \tilde{\sigma}_f^{\bar{k}}(s)$ for all $k \le \bar{k}$ By assumption, we have that $|\Delta x(0)| \le \overline{s}_0$ and $|\Delta u(j)| \le \|\Delta \mathbf{u}\| \le \overline{s}_u$ for all $j \in \mathbb{I}_{\ge 0}$. Thus we have that

$$|\Delta x(k)| \le \tilde{\sigma}_f^{\bar{k}}(\overline{s}_0) \oplus \bigoplus_{j=0}^{\bar{k}-1} \tilde{\sigma}_f^{\bar{k}-j-1}(\overline{s}_u) \coloneqq \overline{s}_x.$$

Furthermore, we have that

$$\begin{aligned} |\Delta y(k)| &\le \sigma_h(|\Delta x(k)|) \\ &\le \sigma_h(\overline{s}_x) \coloneqq \overline{s}_y, \end{aligned}$$

which completes this proof. □

In order to apply this proposition towards the continuity of $V(\cdot)$, we need upper bounds on $|\Delta x|$, $\|\Delta \mathbf{u}^*\|_{0:k-1}$, and k^*. An upper bound on $|\Delta x|$ depends only on where $V(\cdot)$ is evaluated, and thus is straightforward. However, upper bounds on the other two terms require significantly more work. Because we are optimizing over an unbounded domain, there is no guarantee the infimum used in the construction $V(\cdot)$ is attained by any time, pair of input sequences, and parameter sequence. However, any sequences that are "sufficiently close" to optimal admit upper bounds on $\|\Delta \mathbf{u}^*\|_{0:k-1}$ and k^*.

Proposition 14. *Let $V : \mathbb{X} \times \mathbb{X} \to \mathbb{R}_{\ge 0}$ be defined in eq.* (11)*, let $M \ge 1$, and let $\mathcal{B}(M) \coloneqq \{(x_1,x_2) \in \mathbb{X} \times \mathbb{X} : 1/M \le |x_1 - x_2| \le M\}$. There exist $\bar{\varepsilon}(M) > 0$, $\bar{k}(M) \in \mathbb{I}_{\ge 0}$, and $\overline{s}_u(M) > 0$ such that for all $(x_1,x_2) \in \mathcal{B}(M)$ and $\varepsilon \le \bar{\varepsilon}(M)$, if we have $\mathbf{u}_1^*, \mathbf{u}_2^* \in \mathbb{U}^\infty$, $\mathbf{d}^* \in \mathbb{D}^\infty$, and $k^* \in \mathbb{I}_{\ge 0}$ such that*

$$\begin{aligned} V(x_1,x_2) \le \varepsilon + \lambda^{-k^*/2} \Bigg[&\alpha(|\Delta x^*(k^*)|) - \sum_{j=0}^{k^*-1} 2\alpha_u(|\Delta u^*(j)|)\lambda^{k^*-j-1} \\ &- \sum_{j=0}^{k^*-1} \alpha_y(|\Delta y^*(j)|)\lambda^{k^*-j-1} \Bigg], \end{aligned} \tag{15}$$

then $k^* \le \bar{k}(M)$ and $\|\Delta \mathbf{u}^*\|_{0:k^*-1} \le \overline{s}_u(M)$, in which we recall $\lambda \coloneqq e^{-1}$.

Proof. We can substitute eq. (12) into eq. (15) in order to obtain

$$\begin{aligned} V(x_1, x_2) \leq &\varepsilon + \lambda^{-k^*/2} \left[\alpha_x(|\Delta x(0)|)\lambda^{k^*} - \sum_{j=0}^{k^*-1} \alpha_u(|\Delta u^*(j)|)\lambda^{k^*-j-1} \right] \\ \leq &\varepsilon + \alpha_x(|\Delta x(0)|)\lambda^{k^*/2}. \end{aligned}$$

Furthermore, from eq. (14), we have that

$$V(x_1, x_2) \geq \alpha(|\Delta x(0)|).$$

Inspired by [15, Lemma 6.b], define

$$\mu(M) := \max_{s \in [1/M, M]} \frac{\alpha_x(s)}{\alpha(s)}$$

and note that the maximum exists because we are optimizing a continuous function over a compact set. Because of the fashion by which $\alpha_x(\cdot)$ and $\alpha(\cdot)$ are constructed using theorem 7, we have that $\alpha_x(s) \geq \alpha(s)$ for all $s \in \mathbb{R}_{\geq 0}$. Therefore, we have that $\mu(M) \geq 1$, and thus for any $(x_1, x_2) \in \mathcal{B}(M)$, we have that

$$\alpha_x(|\Delta x(0)|) \leq \mu(M)\alpha(|\Delta x(0)|).$$

We can chose $\bar{\varepsilon}(M) := (1/2)\alpha(1/M)$. Fix $(x_1, x_2) \in \mathcal{B}(M)$. We have that

$$\begin{aligned} \alpha(|\Delta x(0)|) \leq V(x_1, x_2) &\leq \varepsilon + \alpha_x(|\Delta x(0)|)\lambda^{k^*/2} \\ &\leq \bar{\varepsilon}(M) + \alpha_x(|\Delta x(0)|)\lambda^{k^*/2} \\ &= (1/2)\alpha(1/M) + \alpha_x(|\Delta x(0)|)\lambda^{k^*/2} \\ &\leq (1/2)\alpha(|\Delta x(0)|) + \mu(M)\alpha(|\Delta x(0)|)\lambda^{k^*/2} \end{aligned}$$

because $1/M \leq |\Delta x(0)| \leq M$. We can rearrange to obtain

$$\begin{aligned} 1/2\alpha(|\Delta x(0)|) &\leq \mu(M)\alpha(|\Delta x(0)|)\lambda^{k^*/2} \\ \frac{1}{2\mu(M)} &\leq \lambda^{k^*/2} \\ \bar{k}(M) := \lceil -2\log_\lambda(2\mu(M)) \rceil &\geq k^* \end{aligned}$$

in which $\lceil \cdot \rceil$ denotes the ceiling function.

Now that we have found $\bar{k}(M)$, we can find $\bar{s}_u(M)$. We have that

$$\alpha(|\Delta x(0)|) \leq V(x_1, x_2) \leq \varepsilon + \lambda^{-k^*/2} \left[\alpha_x(|\Delta x(0)|)\lambda^{k^*} - \sum_{j=0}^{k^*-1} \alpha_u(|\Delta u^*(j)|)\lambda^{k^*-j-1} \right].$$

We can rearrange this expression to obtain

$$\alpha(|\Delta x(0)|) + \lambda^{-k^*/2}\left(\sum_{j=0}^{k^*-1} \alpha_u(|\Delta u^*(j)|)\lambda^{k^*-j-1}\right) \leq \varepsilon + \alpha_x(|\Delta x(0)|)\lambda^{k^*/2}.$$

We can next take individual terms of the summation to obtain

$$\alpha(|\Delta x(0)|) + \alpha_u(|\Delta u^*(j)|)\lambda^{k^*/2-j-1} \leq \varepsilon + \alpha_x(|\Delta x(0)|)\lambda^{k^*/2} \quad \forall j \in \mathbb{I}_{0:k^*-1}.$$

If both $(x_1, x_2) \in \mathcal{B}(M)$ and $\varepsilon \leq \bar{\varepsilon}(M)$, we have that

$$\begin{aligned}&\alpha(|\Delta x(0)|) + \alpha_u(|\Delta u^*(j)|)\lambda^{k^*/2-j-1} \\ &\qquad \leq (1/2)\alpha(1/M) + \mu(M)\alpha(|\Delta x(0)|)\lambda^{k^*/2} \quad \forall j \in \mathbb{I}_{0:k^*-1}.\end{aligned}$$

Because $1/M \leq |\Delta x(0)| \leq M$, we can rearrange to obtain

$$\begin{aligned}\alpha_u(|\Delta u^*(j)|)\lambda^{k^*/2-j-1} &\leq (\mu(M) - 1/2)\alpha(|\Delta x(0)|)\lambda^{k^*/2} \quad \forall j \in \mathbb{I}_{0:k^*-1} \\ |\Delta u^*(j)| &\leq \alpha_u^{-1}\left((\mu(M) - 1/2)\alpha(|\Delta x(0)|)\lambda^{-j-1}\right) \forall j \in \mathbb{I}_{0:k^*-1} \\ &\leq \alpha_u^{-1}\left((\mu(M) - 1/2)\alpha(M)\lambda^{-\bar{k}(M)-1}\right) \forall j \in \mathbb{I}_{0:k^*-1} \\ &:= \overline{s}_u(M) \quad \forall j \in \mathbb{I}_{0:k^*-1}\end{aligned}$$

which concludes the proof. □

Finally, using these supporting propositions, we can show that $V(\cdot)$ is continuous. For $x_1, x_2 \in \mathbb{X}$ not on the diagonal $((x_1, x_2) \in \mathbb{X}^2 \mid x_1 = x_2)$, we can choose M to apply these supporting propositions for all (z_1, z_2) in a neighborhood of (x_1, x_2). It turns out that *all* i-UIOSS Lyapunov functions are continuous on the diagonal, and thus $V(\cdot)$ is continuous everywhere.

Proof of theorem 11. Fix $x_1, x_2 \in \mathbb{X}$ and suppose $x_1 \neq x_2$. For any $\varepsilon > 0$, there exist $\mathbf{u}_1^*$, $\mathbf{u}_2^*$, $\mathbf{d}^*$, and $k^* \in \mathbb{I}_{1:\infty}$ such that

$$\begin{aligned}V(x_1, x_2) \leq \varepsilon + \lambda^{-k^*/2}\Bigg[&\alpha(|\Delta x^*(k^*)|) - \sum_{j=0}^{k^*-1} 2\alpha_u(|\Delta u^*(j)|)\lambda^{k^*-j-1} \\ &- \sum_{j=0}^{k^*-1} \alpha_y(|\Delta y^*(j)|)\lambda^{k^*-j-1}\Bigg]. \end{aligned} \tag{16}$$

In order to use theorems 13 and 14 for all points (z_1, z_2) in a neighborhood of (x_1, x_2), choose $M^* = 2\max(|\Delta x|, 1/|\Delta x|)$. Let $\bar{\varepsilon}(M^*)$, $\bar{k}(M^*)$, and $\overline{s}_u(M^*)$ come from theorem 14. Fix

$\varepsilon \leq \bar{\varepsilon}(M^*)$. For $(z_1, z_2) \in \mathcal{B}(M^*)$, we have that

$$V(z_1, z_2) \geq \lambda^{-k^*/2}\Bigg[\alpha(|\phi(k^*; z_1, \mathbf{u}_1^*, \mathbf{d}^*) - \phi(k^*; z_2, \mathbf{u}_2^*, \mathbf{d}^*)|) - \sum_{j=0}^{k^*-1} 2\alpha_u(|\Delta u^*(j)|)\lambda^{k^*-j-1} - \sum_{j=0}^{k^*-1} \alpha_y(|y(j; z_1, \mathbf{u}_1^*, \mathbf{d}^*) - y(j; z_2, \mathbf{u}_1^*, \mathbf{d}^*)|)\lambda^{k^*-j-1}\Bigg] \tag{17}$$

because $\mathbf{u}_1^*$, $\mathbf{u}_2^*$, $\mathbf{d}^*$, and k^* are feasible in the optimization that produces $V(z_1, z_2)$. We can assume that $(\mathbf{u}_1^*, \mathbf{u}_2^*) \in \mathcal{U}_d(M^*) := \{(\mathbf{u}_1, \mathbf{u}_2) \in \mathbb{U}^\infty \times \mathbb{U}^\infty : \|\mathbf{u}_1 - \mathbf{u}_2\| \leq \overline{s}_u(M^*)\}$ because $\|\mathbf{u}_1^* - \mathbf{u}_2^*\|_{0:k^*-1} \leq \overline{s}_u(M^*)$ by theorem 14, and, by causality, the value function eq. (11) does not depend on any elements of $\mathbf{u}_1^*$ and $\mathbf{u}_2^*$ with $j \geq k^*$. For brevity, let $\bar{k} := \bar{k}(M^*)$ and

$$\begin{aligned} z_1^*(j) &:= \phi(j; z_1, \mathbf{u}_1^*, \mathbf{d}^*) \qquad & \eta_1^*(j) &:= h(z_1^*(j)) \\ z_2^*(j) &:= \phi(j; z_2, \mathbf{u}_2^*, \mathbf{d}^*) \qquad & \eta_2^*(j) &:= h(z_2^*(j)). \end{aligned}$$

We can combine eq. (16) and eq. (17) to obtain

$$V(x_1, x_2) - V(z_1, z_2) \leq \varepsilon + \lambda^{-k^*/2}\Bigg(\alpha(|\Delta x^*(k^*)|) - \alpha(|\Delta z^*(k^*)|) - \Bigg[\sum_{j=0}^{k^*-1} \alpha_y(|\Delta y^*(j)|)\lambda^{k^*-j-1} - \sum_{j=0}^{k^*-1} \alpha_y(|\Delta \eta^*(j)|)\lambda^{k^*-j-1}\Bigg]\Bigg) \tag{18}$$

in which the u terms have been canceled. Let $\Delta V := V(x_1, x_2) - V(z_1, z_2)$. We can take the absolute value of this bound to obtain

$$\begin{aligned} \Delta V \leq& \varepsilon + \lambda^{-k^*/2}\Bigg|\alpha(|\Delta x^*(k^*)|) - \alpha(|\Delta z^*(k^*)|) \\ &- \Bigg[\sum_{j=0}^{k^*-1} \alpha_y(|\Delta y^*(j)|)\lambda^{k^*-j-1} - \sum_{j=0}^{k^*-1} \alpha_y(|\Delta \eta^*(j)|)\lambda^{k^*-j-1}\Bigg]\Bigg| \\ \leq& \varepsilon + \lambda^{-k^*/2}\Bigg(\Big|\alpha(|\Delta x^*(k)|) - \alpha(|\Delta z^*(k)|)\Big| \\ &+ \sum_{j=0}^{k^*-1} \Big|\alpha_y(|\Delta y^*(j)|) - \alpha_y(|\Delta \eta^*(j)|)\Big|\lambda^{k^*-j-1}\Bigg) \end{aligned}$$

Because both $(x_1, x_2), (z_1, z_2) \in \mathcal{B}(M^*)$ and $(\mathbf{u}_1^*, \mathbf{u}_2^*) \in \mathcal{U}_d(M^*)$, by theorem 13 we have that

$$\begin{aligned} |\Delta x^*(k^*)| &\leq \overline{s}_x(M^*) \\ |\Delta z^*(k^*)| &\leq \overline{s}_x(M^*) \\ |\Delta y^*(j)| &\leq \overline{s}_y(M^*) \quad \text{for all } j \leq k^* \\ |\Delta \eta^*(j)| &\leq \overline{s}_y(M^*) \quad \text{for all } j \leq k^*, \end{aligned}$$

in which we have abbreviated $\overline{s}_x(M^*, \overline{s}_u(M^*), \bar{k}(M^*))$ and $\overline{s}_y(M^*, \overline{s}_u(M^*), \bar{k}(M^*))$ as $\overline{s}_x(M^*)$ and $\overline{s}_y(M^*)$, because $k^* \leq \bar{k}(M^*)$. By Proposition 20 in [4], there exist $\sigma, \sigma_y \in \mathcal{K}_\infty$ such that

$$\begin{aligned} |\alpha(s_1) - \alpha(s_2)| &\leq \sigma(|s_1 - s_2|) \quad \text{for all } s_1, s_2 \in [0, \overline{s}_x(M^*)] \\ |\alpha_y(s_1) - \alpha_y(s_2)| &\leq \sigma_y(|s_1 - s_2|) \quad \text{for all } s_1, s_2 \in [0, \overline{s}_y(M^*)]. \end{aligned}$$

We thus have that

$$\begin{aligned} \Delta V \leq& \varepsilon + \lambda^{-k^*/2} \Bigg(\sigma\left(\big| |x_1^*(k^*) - x_2^*(k^*)| - |z_1^*(k^*) - z_2^*(k^*)| \big| \right) \\ &+ \sum_{j=0}^{k^*-1} \sigma_y \left(\big| |y_1^*(j) - y_2^*(j)| - |\eta_1^*(j) - \eta_2^*(j)| \big| \right) \lambda^{k^*-j-1} \Bigg) \\ \leq& \varepsilon + \lambda^{-k^*/2} \Bigg(\sigma\left(|x_1^*(k^*) - z_1^*(k^*)| + |x_2^*(k^*) - z_2^*(k^*)| \right) \\ &+ \sum_{j=0}^{k^*-1} \sigma_y \left(|y_1^*(j) - \eta_1^*(j)| + |y_2^*(j) - \eta_2^*(j)| \right) \lambda^{k^*-j-1} \Bigg) \\ \leq& \varepsilon + \lambda^{-k^*/2} \Bigg(\sigma(2\, |x_1^*(k^*) - z_1^*(k^*)|) + \sigma(2\, |x_2^*(k^*) - z_2^*(k^*)|) \\ &+ \sum_{j=0}^{k^*-1} \Big(\sigma_y(2\, |y_1^*(j) - \eta_1^*(j)|) + \sigma_y(2\, |y_2^*(j) - \eta_2^*(j)|) \Big) \lambda^{k^*-j-1} \Bigg) \end{aligned}$$

in which the states and outputs change partners by applying theorem 12. By theorem 9, there exist $\sigma_f, \sigma_h \in \mathcal{K}$ such that $|f(x_1, u, d) - f(x_2, u, d)| \leq \sigma_f(|\Delta x|)$ and $|h(x_1) - h(x_2)| \leq \sigma_h(|\Delta x|)$ for all $x_1, x_2 \in \mathbb{X}$, $u \in \mathbb{U}$, and $d \in \mathbb{D}$. We can apply these bounds recursively to obtain

$$|\Delta x(j)| \leq \sigma_f^j(|\Delta x(0)|).$$

We thus have that

$$\begin{aligned} \Delta V \leq& \varepsilon + \lambda^{-k^*/2} \Bigg(\sigma(2\sigma_f^{k^*}(|x_1 - z_1|)) + \sigma(2\sigma_f^{k^*}(|x_2 - z_2|)) \\ &+ \sum_{j=0}^{k^*-1} \Big(\sigma_y(2\sigma_h(\sigma_f^j(|x_1 - z_1|))) + \sigma_y(2\sigma_h(\sigma_f^j(|x_2 - z_2|))) \Big) \lambda^{k^*-j-1} \Bigg) \end{aligned} \tag{19}$$

We can thus define

$$\nu(s) := \max_{k \in \mathbb{I}_{0:\bar{k}}} \lambda^{-k/2} \left(\sigma(2\sigma_f^k(s)) + \sum_{j=0}^{k-1} \sigma_y(2\sigma_h(\sigma_f^j(s)))\lambda^{k-j-1} \right)$$

and note that, because the set of $\mathcal{K}$-functions is closed under addition, strictly positive scalar multiplication, and maximization, $\nu \in \mathcal{K}$.

Thus we have that

$$\Delta V \leq \varepsilon + \nu(|x_1 - z_1|) + \nu(|x_2 - z_2|)$$

because $k^* \leq \bar{k}$ for $\varepsilon \leq \bar{\varepsilon}(M^*)$. Since this upper bound no longer depends on k^*, we can take the limit as $\varepsilon \to 0$ in order to obtain

$$\Delta V \leq \nu(|x_1 - z_1|) + \nu(|x_2 - z_2|)$$

for all $(x_1, x_2), (z_1, z_2) \in \mathcal{B}(M^*)$. Finally, this bound applies for $V(z_1, z_2) - V(x_1, x_2)$ by symmetry, and thus

$$|V(x_1, x_2) - V(z_1, z_2)| \leq \nu(|x_1 - z_1|) + \nu(|x_2 - z_2|)$$

implying that $V(\cdot)$ is $\mathcal{K}$-continuous on $\mathcal{B}(M^*)$. Because $\mathcal{B}(M^*)$ contains all $(z_1, z_2) \in \mathbb{X} \times \mathbb{X}$ within a neighborhood of (x_1, x_2), we have that $V(\cdot)$ is continuous at (x_1, x_2).

Now suppose $x_1 = x_2 = x$. Then we have that $V(x_1, x_2) = 0$, and thus

$$\begin{aligned} |V(z_1, z_2) - V(x, x)| = V(z_1, z_2) &\leq \alpha_x(|z_1 - z_2|) \\ &\leq \alpha_x(|z_1 - x| + |x - z_2|) \\ &\leq \alpha_x(2\,|z_1 - x|) + \alpha_x(2\,|z_2 - x|) \end{aligned}$$

and, as a result, $V(\cdot)$ is continuous at (x, x). Thus $V(\cdot)$ is continuous on all of $\mathbb{X} \times \mathbb{X}$. □

4 Necessity of $V(x_1, x_2)$ construction

Note, though we have not explicitly required it, we can assume that $V(\cdot)$ is symmetric as a function of x_1 and x_2, i.e., $V(x_1, x_2) = V(x_2, x_1)$. Indeed, if we have an asymmetric $V(\cdot)$, then we can define $\tilde{V}(x_1, x_2) := (1/2)(V(x_1, x_2) + V(x_2, x_1))$ to be a symmetric i-UIOSS Lyapunov function (Angeli [7] made the same observation for incremental ISS).

We have defined $V(\cdot)$ as a function from $\mathbb{X}^2 \to \mathbb{R}_{\geq 0}$. We might wonder if every i-UIOSS system admits a i-UIOSS of the form $V(x_1, x_2) = \Lambda(x_1 - x_2)$ for $\Lambda : \mathbb{X} \to \mathbb{R}_{\geq 0}$. Angeli [7] provides an example of a continuous-time system that is incrementally exponentially stable which does not admit an incremental Lyapunov function of that form. Because such a system can be interpreted as i-UIOSS that has no dependence on the input u or time-varying parameter d when it is augmented with the trivial output function $h(x) = 0$, that example shows that no such construction is possible in continuous time.

Inspired by the continuous-time system proposed in [7], consider the piecewise-linear system defined by

$$x^+ = f(x) := \begin{cases} A_1 x & \xi_1 \geq 0 \\ A_2 x & \xi_1 < 0 \end{cases}, \quad \text{in which} \quad A_1 = \begin{bmatrix} 1/2 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

and also in which we write $x = [\xi_1 \quad \xi_2]'$ to avoid ambiguity in notation when referring to individual elements of the vector x at the same time as two different states x_1 and x_2.

We have chosen A_1 and A_2 such that $A_1 = A_2 + [1/2 \quad 1]'[1 \quad 0]$, and so, for $\xi_1 = 0$, we have that $A_1 x = A_2 x$ and thus the system is Lipschitz continuous. Note also that A_1 is Schur stable, while A_2 is the matrix that corresponds to a 90° rotation clockwise. Because A_1 has only nonnegative entries, the positive quadrant is forward invariant. In fact, if $\xi_1 \geq 0$ and $\xi_1 \geq -2\xi_2$, the state is mapped to the ray $\xi_1 \geq 0$ and $\xi_2 = 0$. If we have that $\xi_1 \geq 0$ but $\xi_1 \leq -2\xi_2$, the state is mapped to the ray $\xi_1 \leq 0$ and $\xi_2 = 0$. For any x such that $\xi_1 \leq 0$, the state is rotated into the positive quadrant. Thus, every state is mapped to the positive ray $\xi_1 \geq 0$ and $\xi_2 = 0$ in three steps, where ξ_1 exponentially decays.

Because $f(\cdot)$ is Lipschitz continuous, for any $x_1, x_2 \in \mathbb{R}^2$, we have that

$$\left|x_1^+ - x_2^+\right| \leq L \left|x_1 - x_2\right|$$

for a Lipschitz constant $L \approx 1.118 > 1$. We thus have that

$$|x_1(3) - x_2(3)| \leq L^3 \left|x_1(0) - x_2(0)\right| \tag{20}$$

at which point, irrespective of the initial condition, $x_1(3)$ and $x_2(3)$ are on the $\xi_1 = 0$ axis where $f(\cdot)$ is a contraction map. Thus we have that

$$|x_1(k) - x_2(k)| \leq 2^3 L^3 (1/2)^k \left|x_1(0) - x_2(0)\right| \tag{21}$$

and thus this system is incrementally exponentially stable.

Suppose the system admits an incremental Lyapunov function of the form $\Lambda(x_1 - x_2)$. Then $\Lambda(x)$ serves as a Lyapunov function for the stable equilibrium $x = 0$. According to the first paragraph of this section, we can assume $\Lambda(x_1 - x_2) = \Lambda(x_2 - x_1)$, and therefore $\Lambda(x) = \Lambda(-x)$. Thus, because we have

$$\Lambda(A_2 x) \leq \eta \Lambda(x), \tag{22}$$

in which $\eta \in (0, 1)$, for x such that $\xi_1 \leq 0$, we also have

$$\Lambda(A_2(-x)) \leq \eta \Lambda(-x) \tag{23}$$

by symmetry of $\Lambda(\cdot)$, and obtain a descent condition for x such that $\xi_1 \geq 0$. Thus $\Lambda(\cdot)$ would be a Lyapunov function for the system $x^+ = A_2 x$, which is impossible because A_2 is not Schur stable. Therefore this system does not admit an incremental Lyapunov function of the form $\Lambda(x_1 - x_2)$.

5 Changing supply rates

A useful result, first demonstrated in Sontag and Teel [34] in continuous time and in Nesic and Teel [26] in discrete time, is that the the composition of a smooth convex $\mathcal{K}_\infty$ function with an ISS Lyapunov function is itself an ISS Lyapunov function.[1] An extension to nonnegative storage functions is provided by Lemma 4 in Grimm et al. [12]. This extension is general enough that we can apply it directly to i-IOSS Lypaunov functions. However, an additional observation permits us to extend this result to *all* convex functions, not just smooth ones.

Let $\rho(\cdot)$ be a differentiable convex $\mathcal{K}_\infty$ function and let $q(\cdot)$ be its derivative. The key observation that enables these results is that, by the mean value theorem, $\rho(s_0) - \rho(s) \leq q(s_0)(s_0 - s)$ for all $s, s_0 \geq 0$. This inequality can be rearranged to obtain

$$\rho(s) \geq \rho(s_0) + q(s_0)(s - s_0) \tag{24}$$

for all $s, s_0 \geq 0$. We next need the concept of a subgradient from convex analysis (see, e.g., [32, Ch. 23]).

Definition 15. For a convex function $\rho : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, u is a subgradient of $\rho(\cdot)$ at a point s_0 if

$$\rho(s) \geq \rho(s_0) + u(s - s_0)$$

for all $s \in \mathbb{R}_{\geq 0}$. The set of all subgradients at a point s_0 is called the subdifferential of $\rho(\cdot)$ at s_0, and is referred to by $\partial\rho(s_0)$. The subdifferential $\partial\rho : \mathbb{R}_{\geq 0} \rightrightarrows \mathbb{R}$ is a set-valued map. ■

As we would expect, the only subgradient of a differentiable function is its derivative (for points in the interior of its domain). In order to apply eq. (24) for non-differentiable functions, it is sufficient that $q(s) \in \partial\rho(s)$ for all $s \in \mathbb{R}_{\geq 0}$. For convenience, we define $q(s) := \max \partial\rho(s)$, i.e., $q(\cdot)$ is the right derivative of $\rho(\cdot)$. It can be shown that $q(\cdot)$ is a nondecreasing function. Because the continuity of $q(\cdot)$ is never used in the proof of Grimm et al. [12, Lemma 4], we can reproduce that lemma here.

Proposition 16 (Grimm et al. [12, Lemma 4])**.** *Let $W : \mathbb{X} \to \mathbb{R}_{\geq 0}$, be some storage function, $\sigma : \mathbb{X} \to \mathbb{R}_{\geq 0}$ be some measure of state size, and $\ell : \mathbb{X} \times \mathbb{U} \to \mathbb{R}_{\geq 0}$ be some supply rate such that*

$$W(x) \leq \alpha_2(\sigma(x))$$
$$W(f(x,u)) - W(x) \leq \gamma(\ell(x,u)) - \alpha_3(\sigma(x))$$

for all $x \in \mathbb{X}$ and $u \in \mathbb{U}$, in which $\alpha_2, \alpha_3 \in \mathcal{K}_\infty$ and $\gamma \in \mathcal{K}$. Furthermore, let $\rho \in \mathcal{K}_\infty$ be convex, and let $q(\cdot)$ be its right derivative. Then we have that

$$\rho \circ W(f(x,u)) - \rho \circ W(x) \leq 2q \circ \theta(\ell(x,u)) \cdot \gamma(\ell(x,u)) - \frac{1}{4} q\left(\frac{1}{4}\alpha_3(\sigma(x))\right) \alpha_3(\sigma(x))$$

in which $\theta(s) := \gamma(s) + \alpha_2 \circ \alpha_3^{-1} \circ 2\gamma(s)$ is a $\mathcal{K}$ function.

[1] Rather than using the language of "convexity", these references invoked a smooth $\mathcal{K}_\infty$ function whose derivative is nondecreasing. Of course, such a function is convex.

We would like to end up with a $\mathcal{K}_\infty$ function dissipation rate $\tilde{\alpha}_3(\cdot)$ and a $\mathcal{K}$ function supply gain $\tilde{\gamma}(\cdot)$. While the algebra of $\mathcal{K}$ functions is well known, the presence of $q(\cdot)$ complicates matters. The composition of a nondecreasing function and a $\mathcal{K}$ function is nondecreasing, and the product of a nondecreasing function and a $\mathcal{K}$ ($\mathcal{K}_\infty$) function is $\mathcal{K}$ ($\mathcal{K}_\infty$) so long as that nondecreasing function is strictly positive for strictly positive s. Because $q(\cdot)$ is monotone, it is integrable, and we have that

$$\rho(s) = \int_0^s q(t)dt$$

as a result of [32, Theorem 24.2]. Because $\mathcal{K}$ functions are nonnegative, $0 \in \partial\rho(0)$, and thus $q(0) \geq 0$. If $q(s^*) = 0$ for some $s^* > 0$, we would have that $q(s) = 0$ on $[0, s^*]$, and thus $\rho(s^*) = 0$. But because $\rho \in \mathcal{K}_\infty$, that produces a contradiction, and thus $q(s) > 0$ for all $s > 0$.

If $q(\cdot)$ is continuous, then the supply gain and dissipation rate for $\rho \circ W(\cdot)$ in theorem 16 are $\mathcal{K}$ and $\mathcal{K}_\infty$ functions, respectively. If $q(\cdot)$ is not continuous, then we need to find upper and lower bounds for it in terms of continuous monotone functions. Proposition 4 in Rawlings and Risbeck [30] gives us a $\mathcal{K}$ function lower bound for $q(\cdot)$, and, by a similar construction, it is possible to find a continuous upper bound $\tilde{q} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$. As a result, we can finally state a theorem.

Theorem 17. *Let $W : \mathbb{X} \to \mathbb{R}_{\geq 0}$, be some storage function, $\sigma : \mathbb{X} \to \mathbb{R}_{\geq 0}$ be some measure of state size, and $\ell : \mathbb{X} \times \mathbb{U} \to \mathbb{R}_{\geq 0}$ be some supply rate such that*

$$\begin{aligned} W(x) &\leq \alpha_2(\sigma(x)) \\ W(f(x,u)) - W(x) &\leq \gamma(\ell(x,u)) - \alpha_3(\sigma(x)) \end{aligned}$$

for all $x \in \mathbb{X}$ and $u \in \mathbb{U}$, in which $\alpha_2, \alpha_3 \in \mathcal{K}_\infty$ and $\gamma \in \mathcal{K}$. Furthermore, let $\rho \in \mathcal{K}_\infty$ be convex. Then there exist $\tilde{\alpha}_2, \tilde{\alpha}_3 \in \mathcal{K}_\infty$ and $\tilde{\gamma} \in \mathcal{K}$ such that

$$\begin{aligned} \rho \circ W(x) &\leq \tilde{\alpha}_2(\sigma(x)) \\ \rho \circ W(f(x,u)) - \rho \circ W(x) &\leq \tilde{\gamma}(\ell(x,u)) - \tilde{\alpha}_3(\sigma(x)) \end{aligned}$$

for all $x \in \mathbb{X}$ and $u \in \mathbb{U}$, i.e., $\rho \circ W(\cdot)$ is also a storage function.

We can finally relate this result to i-IOSS Lyapunov functions.

Corollary 18. *Let $\Lambda : \mathbb{X} \times \mathbb{X} \to \mathbb{R}_{\geq 0}$ be an i-IOSS Lyapunov function and $\rho(\cdot)$ a convex $\mathcal{K}_\infty$ function. Then $\rho \circ \Lambda(\cdot)$ is also an i-IOSS Lyapunov function.*

Proof. Let $\tilde{\Lambda}(x_1, x_2) \coloneqq \rho \circ \Lambda(x_1, x_2)$. We observe that an i-IOSS Lyapunov function is a storage function for the stacked system

$$\begin{bmatrix} x_1^+ \\ x_2^+ \end{bmatrix} = \begin{bmatrix} f(x_1, u_1, d) \\ f(x_2, u_2, d) \end{bmatrix}$$

with $\sigma(x) = |x_1 - x_2|$, supply gain $\gamma(s) = s$, and supply rate

$$\ell(x,u) = \sigma_y(|h(x_1) - h(x_2)|) + \sigma_u(|u_1 - u_2|).$$

Immediately, we have $\tilde{\alpha}_1(\cdot) \coloneqq \rho \circ \alpha_1(\cdot)$ and $\tilde{\alpha}_2(\cdot) \coloneqq \rho \circ \alpha_2(\cdot)$. For the dissipation condition, we have

$$\begin{aligned}\tilde{\Lambda}(x_1^+, x_2^+) &\leq \tilde{\Lambda}(x_1, x_2) - \tilde{\alpha}_3(|x_1 - x_2|) + \tilde{\gamma}\big(\sigma_y(|h(x_1) - h(x_2)|) + \sigma_u(|u_1 - u_2|)\big) \\ &\leq \tilde{\Lambda}(x_1, x_2) - \tilde{\alpha}_3(|x_1 - x_2|) + \tilde{\gamma} \circ 2\sigma_y(|h(x_1) - h(x_2)|) \\ &\qquad + \tilde{\gamma} \circ 2\sigma_u(|u_1 - u_2|).\end{aligned}$$

and thus $\tilde{\Lambda}(\cdot)$ is an i-IOSS Lyapunov function. □

6 Conclusion

We have proposed a new characterization of i-UIOSS using "convolution-maximization" that reveals the strong properties that systems with ISS-like properties already enjoy (at least for discrete-time systems defined on subsets of $\mathbb{R}^n$). Furthermore, we have shown that *any* system that admits a robustly stable state estimator must be i-UIOSS, making it a necessary condition for detectability. We have provided a converse theorem demonstrating that any i-UIOSS system must admit an i-UIOSS Lyapunov function, and, if that system is $\mathcal{K}$ continuous, that i-UIOSS Lyapunov function must also be (not necessarily $\mathcal{K}$) continuous. Finally, we have provided a result on changing supply rates for general storage functions, including i-UIOSS Lyapunov functions.

It is somewhat unsatisfactory that we do not have a $\mathcal{K}$ continuous i-UIOSS Lyapunov function, but consider again a linear system. For any detectable linear system, there exists a positive-definite matrix P such that $V(x_1, x_2) = (x_1 - x_2)'P(x_1 - x_2)$ constitutes an i-IOSS Lyapunov function. Note that this simple quadratic function is not globally $\mathcal{K}$ continuous. It *is*, however, Lipschitz continuous on all bands $\overline{\mathcal{B}}(M) = \{(x_1, x_2) \in \mathbb{X}^2 : |x_1 - x_2| \leq M\}$. We considered a similar family of bands $\mathcal{B}(M)$, but these bands excluded a strip containing the diagonal $\{(x_1, x_2) \in \mathbb{X}^2 \mid x_1 = x_2\}$ in its interior. The function $V(\cdot)$ constructed is $\mathcal{K}$ continuous on each of these bands. Determining how to extend these bounds to the bands $\overline{\mathcal{B}}(M)$ may be an object of future research. However, for an *exponentially* i-UIOSS system, i.e., a system that satisfies theorem 3 with $\mathcal{KL}$ functions of the form $\beta(s, k) = Cs\eta^k$ for $C > 0$ and $\eta \in (0, 1)$, one can construct an i-UIOSS Lyapunov function that is globally *Lipschitz* using an argument similar to the one used in theorem 11.

We hope that i-UIOSS Lyapunov functions will become useful tools for construction and analysis of nonlinear state estimators. A local function similar to an i-UIOSS Lyapunov function was used by Tsinias [43] in the course of the observer design problem, but he requires the observer to be directly involved in the dissipation condition such that it becomes a descent condition for the error dynamics. Whether an i-UIOSS Lyapunov function can be used to directly construct an observer for the system is a subject for future research. In terms of optimization-based state estimation, an i-IOSS Lyapunov function is used in Allan and Rawlings [3] to construct a Lyapunov-like function for FIE.

7 Appendix

Here, we show that the traditional, asymptotic-gain definition of i-UIOSS given by eq. (7) is equivalent to the one in theorem 3. For ease of reference, we define that property here as well.

Definition 19 (Asymptotic-gain incremental uniform input/output-to-state stability)**.** A system eq. (1) satisfies the asymptotic-gain i-UIOSS property if there exist $\beta \in \mathcal{KL}$ and $\gamma_u, \gamma_y \in \mathcal{K}$ such that

$$|\Delta x(k)| \leq \beta_x(|\Delta x(0)|, k) \oplus \gamma_u(\|\Delta \mathbf{u}\|_{0:k-1}) \oplus \gamma_y(\|\Delta \mathbf{y}\|_{0:k-1})$$

for all $x_1, x_2 \in \mathbb{X}$, all $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{U}^\infty$, and $\mathbf{d} \in \mathbb{D}^\infty$. ■

To our knowledge, there is no simple method to show that theorem 19 implies theorem 3 directly. Because we know that a system that admits an i-UIOSS Lyapunov function must also satisfy theorem 3, we can produce an i-UIOSS Lyapunov function using methods similar to those used in Cai and Teel [10] suitably modified to take insights contained in Grüne and Kellett [13] by Allan and Rawlings [3]. We do not use this method in the body of the paper for two reasons. First, the resulting proof is significantly longer and requires the introduction of the concept of SiUGASMIO. Second, although it produces an i-UIOSS Lyapunov function, to our knowledge there is no way to guarantee this method produces a *continuous* i-UIOSS Lyapunov function in case of unbounded $\mathbb{X}$, $\mathbb{U}$, and $\mathbb{D}$. In the case in which these sets are *compact*, then this method should be able to be suitably modified to produce a *smooth* i-UIOSS function using results contained in, e.g., Kellett and Teel [21, 22].

Definition 20 (SiUGASMIO)**.** A system $x^+ = f(x, u, d)$ with output $y = h(x)$ is strong incremental uniform globally asymptotically stable modulo inputs and outputs (SiUGASMIO) if there exists $\beta \in \mathcal{KL}$ and $\alpha_u, \alpha_y, \phi_u, \phi_y \in \mathcal{K}_\infty$ such that the implications

$$\begin{aligned} &|\Delta x(j)| \geq \alpha_u(|\Delta u(j)|) \oplus \alpha_y(|\Delta y(j)|) \quad \forall j \in \mathbb{I}_{0:k-1} \\ &\quad \implies |\Delta x(j)| \leq \beta(|\Delta x(0)|, j) \quad \forall j \in \mathbb{I}_{0:k} \end{aligned}$$

and

$$\begin{aligned} &|\Delta x(k)| < \alpha_u(|\Delta u(k)|) \oplus \alpha_y(|\Delta y(k)|) \\ &\quad \implies |\Delta x(k+1)| \leq \phi_u(|\Delta u(k)|) \oplus \phi_y(|\Delta y(k)|), \end{aligned}$$

for all $x_1, x_2 \in \mathbb{X}$, $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{U}^\infty$, $\mathbf{d} \in \mathbb{D}^\infty$, and $k \in \mathbb{I}_{\geq 0}$. ■

In order to demonstrate that a system satisfying theorem 19 is also SiUGASMIO, we require an alternative characterization of SiUGASMIO.

Proposition 21 (Alternative characterization of SiUGASMIO)**.** *A system is SiUGASMIO if and only if there exists $\alpha_u, \alpha_y \in \mathcal{K}_\infty$ such that we have that:*

1. *There exists $\nu \in \mathcal{K}$ such that we have the implication*

$$\begin{aligned} |\Delta x(j)| \geq &\alpha_u(|\Delta u(j)|) \oplus \alpha_y(|\Delta y(j)|) && \forall j \in \mathbb{I}_{0:k-1} \\ &\implies |\Delta x(j)| \leq \nu(|\Delta x(0)|) && \forall j \in \mathbb{I}_{0:k} \end{aligned}$$

for all $x_1, x_2 \in \mathbb{X}$, $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{U}^\infty$, $\mathbf{d} \in \mathbb{D}^\infty$, and $k \in \mathbb{I}_{\geq 0}$.

2. *For every $\varepsilon > 0$ and $r > 0$, there exists $J(r, \varepsilon) \in \mathbb{I}_{\geq 0}$ such that for all $x_1, x_2 \in \mathbb{X}$ satisfying $|\Delta x(0)| \leq r$, we have the implication*

$$\begin{aligned} |\Delta x(j)| \geq &\alpha_u(|\Delta u(j)|) \oplus \alpha_y(|\Delta y(j)|) \quad \forall j \in \mathbb{I}_{0:k-1} \\ &\implies |\Delta x(j)| \leq \varepsilon \quad \forall j \in \mathbb{I}_{J:k} \end{aligned}$$

for all $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{U}^\infty$, $\mathbf{d} \in \mathbb{D}^\infty$, and $k \in \mathbb{I}_{\geq 0}$.

3. *There exists $\phi_u, \phi_y \in \mathcal{K}$ such that we have the implication*

$$\begin{aligned} |\Delta x(k)| < &\alpha_u(|\Delta u(k)|) \oplus \alpha_y(|\Delta y(k)|) \\ &\implies |\Delta x(k+1)| \leq \phi_u(\Delta u(k)) \oplus \phi_y(\Delta y(k)) \end{aligned}$$

for all $x_1, x_2 \in \mathbb{X}$, $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{U}^\infty$, $\mathbf{d} \in \mathbb{D}^\infty$, and $k \in \mathbb{I}_{\geq 0}$.

Proof. This proof follows that of Proposition 2.3 in Cai and Teel [10] as modified to add the third condition in Proposition 10 in Allan and Rawlings [2]. □

Now we can show that a system satisfying theorem 19 is SiUGASMIO.

Proposition 22. *If a system eq.* (1) *satisfies theorem 19, it is SiUGASMIO.*

Proof. We aim to show that an i-UIOSS system satisfies the three implications in theorem 21. The first two conditions can be proven in a fashion similar to that used in Lemma 3.6 in Cai and Teel [10]. The main change necessary is that the base case in several instances of proof by induction must be modified because the measurement at time k is included in the IOSS bound in that paper, while it is excluded in the i-UIOSS bound in this paper.

Let $\nu(s) := \beta(s, 0)$, and note that, by evaluation of eq. (7) for $k = 0$, we have that $\nu(s) \geq s$ for all $s \geq 0$.

Claim 23. *There exists some $\alpha_u \in \mathcal{K}_\infty$ such that for all $x_1, x_2 \in \mathbb{X}$, $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{U}^\infty$, and $\mathbf{d} \in \mathbb{D}^\infty$, we have the implication*

$$\begin{aligned} |\Delta x(j)| \geq &\alpha_u(|\Delta u(j)|) && \forall j \in \mathbb{I}_{0:k-1} \\ &\implies |\Delta x(j)| \leq \beta(|\Delta x(0)|, j) \oplus \gamma_y(\|\Delta \mathbf{y}\|_{0:j-1}) \oplus \frac{|\Delta x(0)|}{2} && \forall j \in \mathbb{I}_{0:k}. \end{aligned}$$

Proof. Let $\alpha_u(\cdot)$ be such that $\gamma_u \circ \alpha_u^{-1}(s) \leq \nu^{-1}(s)/2$, which implies that

$$\gamma_u \circ \alpha_u^{-1}(s) \leq \frac{\nu^{-1}(s)}{2} \leq \frac{s}{2} \leq s \tag{25}$$

for all $s \geq 0$. We prove this claim by induction. In the base case, the antecedent is true irrespective of the states and inputs. We thus require that

$$|\Delta x(0)| \leq \beta(|\Delta x(0)|, 0) \oplus \frac{|\Delta x(0)|}{2}$$

for all $x_1, x_2 \in \mathbb{X}$. Because $\beta(s, 0) \geq s$ for all $s \geq 0$, the statement holds.

For the inductive case, we have that

$$|\Delta x(j)| \geq \alpha_u(|\Delta u(j)|) \qquad \forall j \in \mathbb{I}_{0:k} \tag{26}$$

$$|\Delta x(j)| \leq \beta(|\Delta x(0)|, j) \oplus \gamma_y(\|\Delta \mathbf{y}\|_{0:j-1}) \oplus \frac{|\Delta x(0)|}{2} \qquad \forall j \in \mathbb{I}_{0:k}. \tag{27}$$

Equation (26) implies

$$\gamma_u(\|\Delta \mathbf{u}_{0:k}\|) \leq \gamma_u \circ \alpha_u^{-1}(\|\Delta \mathbf{x}\|_{0:k}).$$

From eq. (27) and by noting that $\nu(s) \coloneqq \beta(s, 0)$, we have

$$\|\Delta \mathbf{x}\|_{0:k} \leq \nu(|\Delta x(0)|) \oplus \gamma_y(\|\Delta \mathbf{y}\|_{0:k-1}) \oplus \frac{|\Delta x(0)|}{2}.$$

Combining these two expressions with eq. (7) evaluated at $k + 1$, we obtain

$$\begin{aligned}|\Delta x(k+1)| \leq &\beta(|\Delta x(0)|, k+1) \oplus \gamma_y(\|\Delta \mathbf{y}\|_{0:k}) \\ &\oplus \gamma_u \circ \alpha_u^{-1}\left(\nu(|\Delta x(0)|) \oplus \gamma_y(\|\Delta \mathbf{y}\|_{0:k-1}) \oplus \frac{|\Delta x(0)|}{2}\right).\end{aligned}$$

We can then simplify this inequality with eq. (25) to obtain

$$\begin{aligned}|\Delta x(k+1)| \leq &\beta(|\Delta x(0)|, k+1) \oplus \gamma_y(\|\Delta \mathbf{y}\|_{0:k}) \\ &\oplus \frac{|\Delta x(0)|}{2} \oplus \gamma_y(\|\Delta \mathbf{y}\|_{0:k-1}) \oplus \frac{|\Delta x(0)|}{2} \\ = &\beta(|\Delta x(0)|, k+1) \oplus \gamma_y(\|\Delta \mathbf{y}\|_{0:k}) \oplus \frac{|\Delta x(0)|}{2},\end{aligned}$$

which completes the inductive case. □

Claim 24. *There exists some $\alpha_y \in \mathcal{K}_\infty$ such that for all $x_1, x_2 \in \mathbb{X}$, $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{U}^\infty$, and $\mathbf{d} \in \mathbb{D}^\infty$, we have the implication*

$$\begin{aligned}|\Delta x(j)| \geq &\alpha_u(|\Delta u(j)|) \oplus \alpha_y(|\Delta y(j)|) \qquad \forall j \in \mathbb{I}_{0:k-1} \\ &\implies \gamma_y(\|\Delta \mathbf{y}\|_{0:k-1}) \leq \frac{\Delta x(0)}{2} \\ &\quad \textit{and } |\Delta x(j)| \leq \beta(|\Delta x(0)|, j) \oplus \frac{|\Delta x(0)|}{2} \qquad \forall j \in \mathbb{I}_{0:k}.\end{aligned}$$

Proof. Let $\alpha_y \in \mathcal{K}_\infty$ be such that

$$\gamma_y \circ \alpha_y^{-1}(s) \leq \frac{\nu^{-1}(s)}{2} \leq \frac{s}{2} \leq s \tag{28}$$

for all $s \geq 0$. We prove this claim by induction. The proof of the base case is identical to that in theorem 23. For the inductive case, suppose we have that

$$|\Delta x(j)| \geq \alpha_u(|\Delta u(j)|) \oplus \alpha_y(|\Delta y(j)|) \quad \forall j \in \mathbb{I}_{0:k}$$

and that

$$\gamma_y(\|\Delta \mathbf{y}\|_{0:k-1}) \leq \frac{\Delta x(0)}{2}. \tag{29}$$

As a result, we have that

$$\begin{aligned}
\gamma_y(|\Delta y(k)|) &\leq \gamma_y \circ \alpha_y^{-1}(|\Delta x(k)|) \\
&\leq \gamma_y \circ \alpha_y^{-1}\left(\beta(|\Delta x(0)|, k) \oplus \gamma_y(\|\Delta \mathbf{y}\|_{0:k-1}) \oplus \frac{|\Delta x(0)|}{2}\right) \\
&\leq \gamma_y \circ \alpha_y^{-1}\left(\beta(|\Delta x(0)|, k) \oplus \frac{|\Delta x(0)|}{2} \oplus \frac{|\Delta x(0)|}{2}\right)
\end{aligned}$$

by application of theorem 23 and eq. (29). By definition of $\nu(\cdot)$ and eq. (28), we have that

$$\begin{aligned}
\gamma_y(|\Delta y(k)|) &\leq \gamma_y \circ \alpha_y^{-1}\left(\nu(|\Delta x(0)|) \oplus \frac{|\Delta x(0)|}{2}\right) \\
&\leq \frac{|\Delta x(0)|}{2},
\end{aligned}$$

which completes the first half of the implication. Second, we can apply theorem 23 at time $k+1$ to obtain

$$\begin{aligned}
|\Delta x(k+1)| &\leq \beta(|\Delta x(0)|, k+1) \oplus \gamma_y(\|\Delta \mathbf{y}\|_{0:k}) \oplus \frac{|\Delta x(0)|}{2} \\
&\leq \beta(|\Delta x(0)|, k+1) \oplus \frac{|\Delta x(0)|}{2},
\end{aligned}$$

which completes the proof. □

Claim 25. *For every $\varepsilon > 0$ and $r > 0$, there exists $J(r, \varepsilon) \in \mathbb{I}_{\geq 0}$ such that for all $x_1, x_2 \in \mathbb{X}$ satisfying $|\Delta x(0)| \leq r$, we have the implication*

$$\begin{aligned}
|\Delta x(j)| \geq & \alpha_u(|\Delta u(j)|) \oplus \alpha_y(|\Delta y(j)|) \quad \forall j \in \mathbb{I}_{0:k-1} \\
& \implies |\Delta x(j)| \leq \varepsilon \quad \forall j \in \mathbb{I}_{J:k}
\end{aligned}$$

for all $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{U}^\infty$, $\mathbf{d} \in \mathbb{D}^\infty$, and $k \in \mathbb{I}_{\geq 0}$.

Proof. Fix $\varepsilon, r > 0$. There exists some $T \in \mathbb{I}_{\geq 0}$ such that

$$\beta(r \oplus \varepsilon, T) \leq \varepsilon.$$

Suppose that $|\Delta x(0)| \leq r$. We prove the implication

$$\begin{aligned} |\Delta x(iT + j)| \geq& \alpha_u(|\Delta u(iT + j)|) \oplus \alpha_y(|\Delta y(iT + j)|) \quad \forall iT + j \in \mathbb{I}_{0:k-1} \\ \implies& |\Delta x(iT + j)| \leq \varepsilon \oplus r/2^i \quad \forall iT + j \in \mathbb{I}_{0:k} \end{aligned}$$

for $j \in \mathbb{I}_{\geq 0}$ and $i \in \mathbb{I}_{1:\infty}$ by induction in i.

For the base case of $i = 1$, suppose we have that

$$|\Delta x(j)| \geq \alpha_u(|\Delta u(j)|) \oplus \alpha_y(|\Delta y(j)|) \quad \forall j \in \mathbb{I}_{0:k-1}.$$

for some $k \geq T$. Then we can apply theorem 24 to obtain

$$\begin{aligned} |\Delta x(T + j)| &\leq \beta(|\Delta x(0)|, T + j) \oplus \frac{|\Delta x(0)|}{2} \\ &\leq \beta(r, T) \oplus \frac{|r|}{2} \\ &\leq \varepsilon \oplus \frac{|r|}{2} \end{aligned}$$

for all $T + j \in \mathbb{I}_{0:k}$, which completes the case.

For the inductive case, suppose that, for some fixed $i \in \mathbb{I}_{\geq 0}$, we have that

$$|\Delta x(iT + j)| \leq \varepsilon \oplus \frac{|r|}{2^i} \quad \forall iT + j \in \mathbb{I}_{0:k}$$

for all $j \in \mathbb{I}_{\geq 0}$ such that $iT + j \in \mathbb{I}_{0:k}$ and that

$$|\Delta x(j)| \geq \alpha_u(|\Delta u(j)|) \oplus \alpha_y(|\Delta y(j)|) \quad \forall j \in \mathbb{I}_{0:k-1}.$$

for some $k \geq (i + 1)T$. Then we can apply theorem 24 to obtain

$$\begin{aligned} |\Delta x((i + 1)T + j)| &\leq \beta(|\Delta x(iT + j)|, T) \oplus \frac{|\Delta x(iT + j)|}{2} \\ &\leq \beta(\varepsilon \oplus \frac{|r|}{2^i}, T) \oplus \frac{\varepsilon}{2} \oplus \frac{|r|}{2^{i+1}} \\ &\leq \varepsilon \oplus \frac{\varepsilon}{2} \oplus \frac{|r|}{2^{i+1}} \\ &= \varepsilon \oplus \frac{|r|}{2^{i+1}} \end{aligned}$$

which completes the inductive case. Finally, let $n := 1 \oplus \min\{i \in \mathbb{I}_{\geq 0} \mid r/2^i \leq \varepsilon\}$ and $J := nT$. We then have that

$$|\Delta x(j + J)| \leq \varepsilon \oplus \frac{|r|}{2^n} = \varepsilon$$

for all $j \in \mathbb{I}_{0:k-J}$. Thus the claim is established. $\square$

By application of theorem 24, we have the implication

$$\begin{aligned} |\Delta x(j)| \geq & \alpha_u(|\Delta u(j)|) \oplus \alpha_y(|\Delta y(j)|) & \forall j \in \mathbb{I}_{0:k-1} \\ \implies & |\Delta x(j)| \leq \beta(|\Delta x(0)|, k) \oplus \frac{|\Delta x(0)|}{2} & \forall j \in \mathbb{I}_{0:k}. \end{aligned}$$

Because we have that $\beta(s,k) \leq \nu(s)$ for all $s \geq 0$ and $k \in \mathbb{I}_{\geq 0}$, and furthermore that $\nu(s) \geq s/2$ for all $s \geq 0$, we immediately have the first condition of theorem 21. The second condition is given by theorem 25.

Finally, if we have that

$$|\Delta x(k)| < \alpha_u(|\Delta u(k)|) \oplus \alpha_y(|\Delta y(k)|)$$

we can apply eq. (7) to obtain

$$\begin{aligned} |\Delta x(k+1)| &\leq \beta(|\Delta x(k)|, 1) \oplus \gamma_u(|\Delta u(k)|) \oplus \gamma_y(|\Delta y(k)|) \\ &\leq \beta(\alpha_u(|\Delta u(k)|) \oplus \alpha_y(|\Delta y(k)|), 1) \oplus \gamma_u(|\Delta u(k)|) \oplus \gamma_y(|\Delta y(k)|) \\ &= \beta(\alpha_u(|\Delta u(k)|), 1) \oplus \beta(\alpha_y(|\Delta y(k)|), 1) \oplus \gamma_u(|\Delta u(k)|) \oplus \gamma_y(|\Delta y(k)|) \end{aligned}$$

which is a bound of the form required for the third condition to hold. Thus i-UIOSS implies SiUGASMIO. □

In order to produce an i-UIOSS Lyapunov function, we first define an autonomous difference inclusion related to the original system. We obtain a Lyapunov function for that system, then show that it is an i-UIOSS Lyapunov function for the original system.

We first require a converse Lyapunov theorem for difference inclusions with no regularity conditions. For an autonomous difference inclusion

$$x^+ \in F(x)$$

let $S(x)$ denote the set of trajectories originating from x.

Definition 26 (Lyapunov function)**.** A function $V : \mathbb{X} \to \mathbb{R}_{\geq 0}$ is an exponential-decrease Lyapunov function for a difference inclusion $x^+ \in F(x)$ and the (not necessarily closed) set $\mathcal{A}$ if there exist $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$ and $\eta \in (0,1)$ such that

$$\alpha_1(|x|_{\mathcal{A}}) \leq V(x) \leq \alpha_2(|x|_{\mathcal{A}}) \tag{30}$$

$$\sup_{x^+ \in F(x)} V(x^+) \leq \eta V(x) \tag{31}$$

for all $x \in \mathbb{X}$. ■

Proposition 27. *The set $\mathcal{A} \subseteq \mathbb{X}$ is globally asymptotically stable for the difference inclusion $x^+ \in F(x)$, i.e., there exists $\beta \in \mathcal{KL}$ such that*

$$|x(k)|_{\mathcal{A}} \leq \beta(|x(0)|_{\mathcal{A}}, k) \tag{32}$$

along all solutions to the inclusion, if and only if it admits an exponential-decrease Lyapunov function $V(\cdot)$ for that set.

Proof. The proof that such a Lyapunov function is sufficient for $\mathcal{KL}$ stability is completely standard, and so we do not exclude it here. For the necessity of such a Lypaunov function, we note that the argument used in Kellett and Teel [22, Sec. 6.1.1] does not depend on any regularity conditions of the difference inclusion used there, and thus suffices for proof of this proposition. □

Theorem 28. *Every i-UIOSS system $x^+ = f(x, u, d)$ with measurement $y = h(x)$ admits an i-UIOSS Lyapunov function.*

Proof. Because the system is i-UIOSS, it is also SiUGASMIO. Let $\alpha_u, \alpha_y, \phi_u, \phi_y \in \mathcal{K}_\infty$ and $\beta \in \mathcal{KL}$ come from the definition of SiUGASMIO. Define the difference inclusion for a stacked system in which the second copy of the system pads the input u with a ball with radius proportional to $|x_1 - x_2|$:

$$F(x_1, x_2, u, d) \coloneqq \begin{cases} \begin{bmatrix} f(x_1, u, d) \\ f(x_2, (u + (\alpha_u^{-1}(|x_1 - x_2|)\mathbb{B}) \cap \mathbb{U}, d)) \end{bmatrix} & \text{if } |x_1 - x_2| \geq \alpha_y(|h(x_1) - h(x_2)|) \\ \begin{bmatrix} x_1 \\ x_1 \end{bmatrix} & \text{otherwise} \end{cases}$$

in which $\mathbb{B}$ is the closed unit ball. Then, define a map $G(x_1, x_2) \coloneqq F(x_1, x_2, \mathbb{U}, \mathbb{D})$, in which $F(\cdot)$ is evaluated at all possible values of inputs and outputs simultaneously.

Note that the system defined by

$$\begin{bmatrix} x_1^+ \\ x_2^+ \end{bmatrix} \in G(x_1, x_2)$$

is an autonomous difference inclusion. We next show that the set $\mathcal{A} \coloneqq \{(x_1, x_2) \mid x_1 = x_2\}$ is asymptotically stable for this system. As noted in Angeli [6, Lemma 2.3], the functions $|(x_1, x_2)|_{\mathcal{A}}$ and $|x_1 - x_2|$ are equivalent in the sense that

$$|(x_1, x_2)|_{\mathcal{A}} = \frac{\sqrt{2}}{2} |x_1 - x_2| .$$

First, we show that $\mathcal{A}$ is forward invariant. Suppose that $x_1 = x_2$. Then the system evolves according to the first case. Furthermore, the input of the x_2 component is not padded because $x_1 = x_2$. Thus $x_1^+ = x_2^+ = f(x_1, u, d)$ for all $u \in \mathbb{U}$ and $d \in \mathbb{D}$. Because $G(x_1, x_2)$ is the union of $F(x_1, x_2, u, d)$ over all $u \in \mathbb{U}$ and $d \in \mathbb{D}$, we have that $(x_1^+, x_2^+) \in \mathcal{A}$ for all $(x_1^+, x_2^+) \in G(x_1, x_2)$, and thus $\mathcal{A}$ is forward invariant.

Let $z \coloneqq (x_1, x_2)$, and let $\mathbf{z} \in S(z)$. Take some element $z(k)$ from $\mathbf{z}$. Suppose first that $|x_1(j) - x_2(j)| \geq \alpha_y(|h(x_1(j)) - h(x_2(j))|)$ for all $j \in \mathbb{I}_{0:k-1}$. Then the inclusion evolves according to the first case for all those j. In particular, we have that $x_1^+ = f(x_1, u, d)$ and $x_2^+ = f(x_2, u + \Delta u, d)$, in which $|\Delta u| \leq \alpha_u^{-1}(|x_1 - x_2|)$. Because $f(\cdot)$ is SiUGASMIO, we thus have that

$$|x_1(j) - x_2(j)| \leq \beta(|x_1(0) - x_2(0)|, j)$$

for all $j \in \mathbb{I}_{0:k}$. We thus have a $\mathcal{KL}$ function upper bound for $|z(k)|_{\mathcal{A}}$. Now suppose that $|x_1(j^*) - x_2(j^*)| < \alpha_y(|h(x_1(j^*)) - h(x_2(j^*))|)$ for some $j^* \in \mathbb{I}_{0:k-1}$. We then have that

$x_1(j^*+1) = x_2(j^*+1) = x_1(j^*)$, which gives a point $z(j^*+1) \in \mathcal{A}$. Because $\mathcal{A}$ is forward invariant, we have that

$$|z(j)|_{\mathcal{A}} = 0$$

for all $j \in \mathbb{I}_{j^*+1:k}$. Thus the set $\mathcal{A}$ is globally asymptotically stable for the difference inclusion $z^+ \in G(z)$, and thus that inclusion admits a Lyapunov function $V : \mathbb{X}\times\mathbb{X} \to \mathbb{R}_{\geq 0}$.

There exist $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$ such that

$$\alpha_1(|x_1 - x_2|) \leq V(x_1, x_2) \leq \alpha_2(|x_1 - x_2|)$$

for all $x_1, x_2 \in \mathbb{X}$ by the equivalence of $|(x_1, x_2)|_{\mathcal{A}}$ and $|x_1 - x_2|$. By construction of $G(\cdot)$, there exists $\eta \in (0, 1)$ such that whenever we have $|\Delta x| \geq \alpha_y(|\Delta y| \oplus \alpha_u(|\Delta u|)$, we have that

$$V(f(x_1, u_1, d), f(x_2, u_2, d)) \leq \eta V(x_1, x_2)$$

for all $x_1, x_2 \in \mathbb{X}$, $u_1, u_2 \in \mathbb{U}$, and $d \in \mathbb{D}$. Now suppose that $|\Delta x| < \alpha_y(|\Delta y| \oplus \alpha_u(|\Delta u|)$. Because $f(\cdot)$ is SiUGASMIO, we have that

$$\left|\Delta x^+\right| \leq \phi_u(|\Delta u|) \oplus \phi_y(|\Delta y|).$$

We thus have that

$$V(x_1^+, x_2^+) \leq \alpha_2(\left|\Delta x^+\right|) \leq \alpha_2 \circ \phi_u(|\Delta u|) \oplus \alpha_2 \circ \phi_y(|\Delta y|).$$

Let $\sigma_u(s) \coloneqq \alpha_2 \circ \phi_u(s)$ and $\sigma_y(s) \coloneqq \alpha_2 \circ \phi_y(s)$. Irrespective of whether or not $|\Delta x| \geq \alpha_y(|\Delta y| \oplus \alpha_u(|\Delta u|)$, we have that

$$V(x_1^+, x_2^+) \leq \eta V(x_1, x_2) + \sigma_u(|u_1 - u_2|) + \sigma_y(|y_1 - y_2|)$$

because $\eta V(x_1, x_2)$ is nonnegative. Thus $V(\cdot)$ is an (exponential-decrease) i-UIOSS Lyapunov function for the original system $x^+ = f(x, u, d)$. $\square$

Having obtained an i-UIOSS Lyapunov function for a system satisfying theorem 19, we can apply theorem 8 to note that it is i-UIOSS, which completes the proof.

References

[1] D. Aeyels. Generic Observability of Differentiable Systems. *SIAM J. Cont. Opt.*, 19 (5):595–603, 1981. doi: 10.1137/0319037.

[2] D. A. Allan and J. B. Rawlings. An input/output-to-state stability converse theorem for closed positive invariant sets. Technical Report 2018–01, TWCCC Technical Report, December 2018. URL `https://sites.engineering.ucsb.edu/~jbraw/jbrweb-archives/tech-reports/twccc-2018-01.pdf`.

[3] D. A. Allan and J. B. Rawlings. A Lyapunov-like function for full information estimation. In *American Control Conference*, pages 4497–4502, Philadelphia, PA, July 10–12, 2019.

[4] D. A. Allan, C. N. Bates, M. J. Risbeck, and J. B. Rawlings. On the inherent robustness of optimal and suboptimal nonlinear MPC. *Systems & Control Letters*, 106:68 – 78, 2017. ISSN 0167-6911. doi: https://doi.org/10.1016/j.sysconle.2017.03.005. URL `http://www.sciencedirect.com/science/article/pii/S016769111730049X`.

[5] B. D. O. Anderson and J. B. Moore. Detectability and stabilizability of time-varying discrete-time linear systems. *SIAM Journal on Control and Optimization*, 19(1):20–32, 1981.

[6] D. Angeli. A Lyapunov approach to incremental stability properties. *IEEE Transactions on Automatic Control*, 47(3):410–421, 2002.

[7] D. Angeli. Further results on incremental input-to-state stability. *IEEE Transactions on Automatic Control*, 54(6):1386–1391, June 2009. ISSN 0018-9286. doi: 10.1109/TAC.2009.2015561.

[8] F. Bayer, M. Bürger, and F. Allgöwer. Discrete-time Incremental ISS: A framework for Robust NMPC. In *2013 European Control Conference (ECC)*, pages 2068–2073, 2013.

[9] R. W. Brockett. System Theory on Group Manifolds and Coset Spaces. *SIAM J. Cont.*, 10(2):265–284, 1972. doi: 10.1137/0310021.

[10] C. Cai and A. R. Teel. Input–output-to-state stability for discrete-time systems. *Automatica*, 44(2):326 – 336, 2008. ISSN 0005-1098. doi: http://dx.doi.org/10.1016/j.automatica.2007.05.022. http://www.sciencedirect.com/science/article/pii/S0005109807003056.

[11] R. A. Freeman and P. V. Kokotovic. *Robust Nonlinear Control Design: State-space and Lyapunov Techniques*. Birkhäuser, Boston, 1996.

[12] G. Grimm, M. J. Messina, S. E. Tuna, and A. R. Teel. Model predictive control: For want of a local control Lyapunov function, all is not lost. *IEEE Transactions on Automatic Control*, 50(5):546–558, 2005.

[13] L. Grüne and C. M. Kellett. ISS-Lyapunov functions for discontinuous discrete-time systems. *IEEE Transactions on Automatic Control*, 59(11):3098–3103, Nov 2014. ISSN 0018-9286. doi: 10.1109/TAC.2014.2321667.

[14] R. Hermann and A. J. Krener. Nonlinear controllability and observability. *IEEE Transactions on Automatic Control*, 22(5):728–740, 1977.

[15] W. Hu. Robust Stability of Optimization-based State Estimation Under Bounded Disturbances. *ArXiv e-prints*, February 2017.

[16] Z.-P. Jiang and Y. Wang. Input-to-state stability for discrete-time nonlinear systems. *Automatica*, 37:857–869, 2001.

[17] T. Kailath. A view of three decades of linear filtering theory. *IEEE Transactions on Information Theory*, IT-20(2):146–181, March 1974.

[18] R. E. Kalman. Contributions to the theory of optimal control. *Bull. Soc. Math. Mex.*, 5:102–119, 1960.

[19] R. E. Kalman and R. S. Bucy. New results in linear filtering and prediction theory. *Trans. ASME, J. Basic Engineering*, pages 95–108, March 1961.

[20] C. M. Kellett and A. R. Teel. Discrete-time asymptotic controllability implies smooth control-Lyapunov function. *Systems & Control Letters*, 52:349–359, 2004.

[21] C. M. Kellett and A. R. Teel. Smooth Lyapunov functions and robustness of stability for difference inclusions. *Systems & Control Letters*, 52:395–405, 2004.

[22] C. M. Kellett and A. R. Teel. On the robustness of $\mathcal{KL}$-stability for difference inclusions: Smooth discrete-time Lyapunov functions. *SIAM Journal on Control and Optimization*, 44(3):777–800, 2005. doi: 10.1137/S0363012903435862. URL `https://doi.org/10.1137/S0363012903435862`.

[23] S. Knüfer and M. A. Müller. Robust Global Exponential Stability for Moving Horizon Estimation. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 3477–3482, Dec 2018. doi: 10.1109/CDC.2018.8619617.

[24] M. Krichman, E. D. Sontag, and Y. Wang. Input-output-to-state stability. *SIAM Journal on Control and Optimization*, 39(6):1874–1928, 2001.

[25] M. A. Müller. Nonlinear moving horizon estimation in the presence of bounded disturbances. *Automatica*, 79:306–314, 2017. ISSN 0005-1098. doi: https://doi.org/10.1016/j.automatica.2017.01.033. URL `http://www.sciencedirect.com/science/article/pii/S0005109817300432`.

[26] D. Nesic and A. R. Teel. Changing supply functions in input to state stable systems: the discrete-time case. *IEEE Transactions on Automatic Control*, 46(6):960–962, June 2001. ISSN 0018-9286. doi: 10.1109/9.928607.

[27] H. Nijmeijer. Observability of autonomous discrete time nonlinear systems: A geometric approach. *International Journal of Control*, 36(5):867–874, 1982.

[28] A. Pavlov and L. Marconi. Incremental passivity and output regulation. *Systems & Control Letters*, 57(5):400–409, 2008. ISSN 0167-6911. doi: https://doi.org/10.1016/j.sysconle.2007.10.008. URL `http://www.sciencedirect.com/science/article/pii/S0167691107001417`.

[29] J. B. Rawlings and D. Q. Mayne. *Model Predictive Control: Theory and Design*. Nob Hill Publishing, Madison, WI, 2009. 668 pages, ISBN 978-0-9759377-0-9.

[30] J. B. Rawlings and M. J. Risbeck. On the equivalence between statements with epsilon-delta and K-functions. Technical Report 2015–01, TWCCC Technical Report, December 2015. URL `http://jbrwww.che.wisc.edu/tech-reports/twccc-2015-01.pdf`.

[31] J. B. Rawlings, D. Q. Mayne, and M. M. Diehl. *Model Predictive Control: Theory, Design, and Computation*. Nob Hill Publishing, Madison, WI, 2nd edition, 2017. 770 pages, ISBN 978-0-9759377-3-0.

[32] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, N.J., 1970.

[33] H. Santoso, D. Hioe, J. Bao, and P. L. Lee. Operability analysis of nonlinear processes based on incremental dissipativity. *J. Proc. Cont.*, 22(1):156–166, Jan 2012. ISSN 0959-1524. doi: 10.1016/J.JPROCONT.2011.09.006. URL `https://www.sciencedirect.com/science/article/pii/S0959152411001910?via{\%}3Dihub`.

[34] E. Sontag and A. Teel. Changing supply functions in input/state stable systems. *IEEE Transactions on Automatic Control*, 40(8):1476–1478, Aug 1995. ISSN 0018-9286. doi: 10.1109/9.402246.

[35] E. D. Sontag. On the Observability of Polynomial Systems, I: Finite-Time Problems. *SIAM J. Cont. Opt.*, 17(1):139–151, jan 1979. ISSN 0363-0129. doi: 10.1137/0317011. URL `https://doi.org/10.1137/0317011`.

[36] E. D. Sontag. A concept of local observability. *Systems & Control Letters*, 5(1):41–47, 1984. ISSN 0167-6911. doi: https://doi.org/10.1016/0167-6911(84)90007-0. URL `http://www.sciencedirect.com/science/article/pii/0167691184900070`.

[37] E. D. Sontag. Smooth stabilization implies coprime factorization. *IEEE Transactions on Automatic Control*, 34(4):435–443, Apr 1989. ISSN 0018-9286. doi: 10.1109/9.28018.

[38] E. D. Sontag. Comments on integral variants of ISS. *Systems & Control Letters*, 34:93–100, 1998.

[39] E. D. Sontag and Y. Wang. On the characterization of the input to state stability property. *Systems & Control Letters*, 24:351–359, 1995.

[40] E. D. Sontag and Y. Wang. Output-to-state stability and detectability of nonlinear systems. *Systems & Control Letters*, 29:279–290, 1997.

[41] H. J. Sussmann. Existence and uniqueness of minimal realizations of nonlinear systems. *Mathematical systems theory*, 10(1):263–284, 1976. ISSN 1433-0490. doi: 10.1007/BF01683278. URL `https://doi.org/10.1007/BF01683278`.

[42] D. N. Tran, B. S. Rüffer, and C. M. Kellett. Incremental stability properties for discrete-time systems. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 477–482, 2016.

[43] J. Tsinias. Further results on the observer design problem. *Systems & Control Letters*, 14(5):411–418, 1990. ISSN 0167-6911. doi: https://doi.org/10.1016/0167-6911(90)90092-9. URL `http://www.sciencedirect.com/science/article/pii/0167691190900929`.

[44] W. M. Wonham. On pole assignment in multi-input controllable linear systems. *IEEE Trans. Auto. Cont.*, 12(6):660–665, 1967.