Interconnected Observers for Robust Decentralized Estimation with Performance Guarantees and Optimized Connectivity Graph

Yuchun Li and Ricardo G. Sanfelice

***Abstract*—Motivated by the need of observers that are both robust to disturbances and guarantee fast convergence to zero of the estimation error, we propose an observer for linear time-invariant systems with noisy output that consists of the combination of N coupled observers over a connectivity graph. At each node of the graph, the output of these *interconnected observers* is defined as the average of the estimates obtained using local information. The convergence rate and the robustness to measurement noise of the proposed observer's output are characterized in terms of $\mathcal{KL}$ bounds. Several optimization problems are formulated to design the proposed observer so as to satisfy a given rate of convergence specification while minimizing the H_∞ gain from noise to estimates or the size of the connectivity graph. It is shown that that the interconnected observers relax the well-known tradeoff between rate of convergence and noise amplification, which is a property attributed to the proposed innovation term that, over the graph, couples the estimates between the individual observers. Sufficient conditions involving information of the plant only, assuring that the estimate obtained at each node of the graph outperforms the one obtained with a single, standard Luenberger observer are given. The results are illustrated in several examples throughout the paper.**

I. Introduction

We consider linear time-invariant systems of the form

$$\dot{x} = Ax, \quad y = Cx + m(t), \tag{1}$$

where $x \in \mathbb{R}^n$, $y \in \mathbb{R}^p$, and $t \mapsto m(t)$ denotes measurement noise, for which there exists a Luenberger observer

$$\dot{\hat{x}}_L = A\hat{x}_L - K_L(\hat{y}_L - y), \quad \hat{y}_L = C\hat{x}_L \tag{2}$$

leading to the exponentially stable error system

$$\dot{e}_L = (A - K_LC)e_L + K_Lm(t) := \tilde{A}_Le_L + K_Lm(t) \tag{3}$$

with estimation error given by $e_L := \hat{x}_L - x$. It is well-known that, under observability conditions of (1), the matrix gain K_L can be chosen to make the convergence rate of (3) arbitrarily fast. However, due to the fast convergence speed requiring a large gain, the price to pay is that the effect of measurement noise m is amplified. Indeed, the design of observers, such as those in the form (2), involves a tradeoff between convergence rate and robustness to measurement noise [1], [2]. In fact, in [1, page 597], D. G. Luenberger makes the following remark about the error system (3) when (C, A) is observable: "Theoretically, the eigenvalues can be moved arbitrarily toward minus infinity, yielding extremely rapid convergence. This tends, however, to make the observer act like a differentiator and thereby become highly sensitive to noise, and to introduce other difficulties." Along the same lines, the authors of [2] recognize the compromise between performance and robustness in the design of (2): "At this point we can only offer some intuitive guidelines for a choice of K to obtain satisfactory performance of the observer. To obtain fast convergence of the reconstruction error to zero, K should be chosen so that the observer poles are quite deep in the left-half complex plane. This, however, generally must be achieved by making the gain matrix K large, which in turn makes the observer very sensitive to any observation noise that may be present, added to the observed variable $y(t)$. A compromise must be found," see [2, page 332]. Unfortunately, this issue is also at the core of every estimation algorithm for multi-agent systems.

Y. Li and R. G. Sanfelice are with the Department of Computer Engineering, University of California, Santa Cruz, CA, 95064, USA. E-mail: yuchunli,ricardo@ucsc.edu. This research has been partially supported by the National Science Foundation under CAREER Grant no. ECS-1150306 and by the Air Force Office of Scientific Research under YIP Grant no. FA9550-12-1-0366.

A. Related work

Several observer architectures and design methods with the goal of conferring good performance and robustness to the error system have been proposed in the literature. In particular, H_∞ tools have been employed to formulate sets of Linear Matrix Inequalities (LMIs) that, when feasible, guarantee that the $\mathcal{L}_2$ gain from disturbance to the estimation error is below a pre-established upper bound; see, e.g., [3], [4], [5], to just list a few. Following ideas from adaptive control [6], observers with a gain that adapts to the plant measurements have been proposed in [7], [8], though the presence of measurement noise can lead to large values of the gains. Such issues also emerge in the design of high-gain observers, where the use of high gain can significantly amplify the effect of measurement noise, as in [9]. More recently, observers using adaptive gains [10], two gains designed with different objectives [11], [12], and switching between two observers [13] have been found successful in certain settings.

Recent research efforts in multi-agent systems have lead to enlightening results in distributed estimation and consensus. Distributed Kalman filtering are employed for achieving spatially-distributed estimation tasks in [14] and for sensor networks in [15], [16], [17], [18], [19]. To characterize the effect of unmodeled dynamics on the consensus multi-agent system, in [20], a region-based approach is used for distributed H_∞-based consensus of multi-agent systems with an undirected graph. For dynamic average consensus, [21] proposes a decentralized algorithm that guarantees asymptotic agreement of a signal over strongly connected and weight-balanced graphs. In [22], switching inter-agent topologies are used to design distributed observers for a leader-follower problem in multi-agent systems. For estimating the trajectory of a moving target with perturbed dynamics, nonlinear filters based on networked sensors are proposed in [23], [24]. However, distributed estimation algorithms that systematically meet specifications of performance and robustness to measurement noise are not available.

B. Contributions

We propose a new observer, called *interconnected observers*, with improved convergence rate of the estimation error and robustness to measurement noise, when compared with the observer in (2). The proposed observer consists of N linear time-invariant observers interconnected over a graph. The local estimate at each node is provided by an observer featuring an innovation term that appropriately injects the estimate obtained from its neighbors, which can be computed in a decentralized manner. The global estimate of the state of the plant is given by the average of the local estimates.

The main contributions of this paper are threefold.

1) We establish that, under certain conditions involving its parameters, and when compared to the Luenberger observer in (2), the proposed observer significantly improves the rate of convergence and the gain from measurement noise to estimation error, with improvements of more than 50% at times (see Table III).
2) We characterize the convergence rate and the robustness to measurement noise of the proposed observer in terms of $\mathcal{KL}$ bounds, which provide useful information on how the parameters of the observers affect such properties.
3) We formulate optimization problems for the purpose of the design of interconnected observers.
    i) For a fixed rate of convergence, optimization problems are proposed for the design of interconnected observers with optimized gain from measurement noise to estimation error (local and global).
    ii) For a fixed rate of convergence and a desired H_∞ gain, optimization problems that minimize the number of edges of the connectivity graph are also formulated.
    iii) An LMI condition involving only information about the plant is provided to guarantee that the estimate obtained at each node of the graph outperforms the one obtained with a single, standard Luenberger observer.

Examples throughout the paper illustrate our results and their applicability to estimation in multi-agent systems, such as mobile and sensor networks. To the best of our knowledge, we are not aware of an observer in the literature that guarantees such properties simultaneously.

C. Organization of the Paper

The remainder of this paper is organized as follows. In Section II, the idea and benefits behind interconnected observers are presented in a motivational example. Section III introduces the proposed observer in general form, the $\mathcal{KL}$ bounds, and the design methods in terms of optimization problems.

II. Motivational Example

Consider the scalar plant

$$\dot{x} = ax, \quad y = x + m, \tag{4}$$

where m denotes measurement noise and $a < 0$. Suppose we want to estimate the state x from measurements of y. Following (2), a Luenberger observer for (4) is given by

$$\dot{\hat{x}}_L = a\hat{x}_L - K_L(\hat{y}_L - y), \quad \hat{y}_L = \hat{x}_L. \tag{5}$$

The resulting estimation error system is given by (3) with $\tilde{A}_L = a - K_L$. Its rate of convergence is $a - K_L$ and, when m is constant, its steady-state error is $e_L^\star := \frac{K_L}{K_L - a} m$. It is apparent that to get fast convergence rate, the constant K_L needs to be positive and large. However, as argued in the introduction, with K_L large, the effect of measurement noise is amplified. In light of recent popularity of multi-agent systems, it is natural to explore the advantages of using more than one measurement of the plant's output so as to overcome such a tradeoff.

In this paper, we show that it is possible to design interconnected observers that are capable of relaxing the said tradeoff. To illustrate the idea behind the proposed observer, consider the estimation of the state of the scalar plant (4) with two agents, each taking its own measurement of y. The two agents can communicate with each other according to the following directed graph: agent 1 can transmit information to agent 2, but agent 2 cannot send data to agent 1. This is shown in Figure 1.

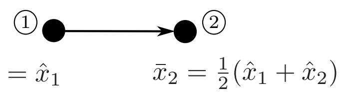


Fig. 1. Two agents connected as a direct graph.

Following the approach in this paper, an interconnected observer would take the form

$$\begin{aligned} &\dot{\hat{x}}_1 = a\hat{x}_1 - K_{11}(\hat{y}_1 - y_1), \\ &\dot{\hat{x}}_2 = a\hat{x}_2 - K_{22}(\hat{y}_2 - y_2) - K_{21}(\hat{y}_1 - y_1), \\ &\hat{y}_1 = \hat{x}_1,\ \hat{y}_2 = \hat{x}_2,\ \bar{x}_1 = \hat{x}_1,\ \bar{x}_2 = (1/2)(\hat{x}_1 + \hat{x}_2), \end{aligned} \tag{6}$$

where $\hat{x}_i$ and $\bar{x}_i$ are associated with agent i, each measured plant output y_i is corrupted by measurement noise m_i, that is $y_1 = x + m_1$ and $y_2 = x + m_2$, respectively, where m_i's are independent. The term "$-K_{21}(\hat{y}_1 - y_1)$" defines an innovation term exploiting the information shared by agent 1 with agent 2. The output $\bar{x}_i$ of agent i defines the local estimate (at agent i) of x. Since agent 1 only has access to its own information, we have $\bar{x}_1 = \hat{x}_1$, while since agent 2 has also information from its neighbor, agent 2's output $\bar{x}_2$ can be taken as the average of the states $\hat{x}_1$ and $\hat{x}_2$.[1]

To analyze the estimation error induced by the interconnected observer in (6), define error variables $e_i := \hat{x}_i - x, i \in \{1, 2\}$. Then, the error system is given by

$$\begin{aligned} &\dot{e}_1 = (a - K_{11})e_1 + K_{11}m_1, \\ &\dot{e}_2 = -K_{21}e_1 + (a - K_{22})e_2 + K_{21}m_1 + K_{22}m_2, \end{aligned} \tag{7}$$

which can be written in matrix form as

$$\dot{e} = \tilde{A}e + \tilde{K}m, \tag{8}$$

where $e = [e_1\ e_2]^\top$, $m = [m_1\ m_2]^\top$,

$$\tilde{A} = \begin{bmatrix} a - K_{11} & 0 \\ -K_{21} & a - K_{22} \end{bmatrix}, \tilde{K} = \begin{bmatrix} K_{11} & 0 \\ K_{21} & K_{22} \end{bmatrix}. \tag{9}$$

Then, when K_{11}, K_{21}, and K_{22} are chosen such that $\tilde{A}$ is Hurwitz and when m is constant, the steady-state value of (8) is given by

$$e_1^\star = \frac{K_{11}}{K_{11} - a} m_1, e_2^\star = \frac{-aK_{21}}{(K_{11} - a)(K_{22} - a)} m_1 + \frac{K_{22}}{K_{22} - a} m_2. \tag{10}$$

Furthermore, the local estimation error resulting from each agent is given by the quantity $\bar{e}_i := \bar{x}_i - x,\ i \in \{1, 2\}$, and has a steady-state value given by

$$\bar{e}_1^\star = e_1^\star,\ \bar{e}_2^\star = \frac{K_{11}(K_{22} - a) - aK_{21}}{2(K_{11} - a)(K_{22} - a)} m_1 + \frac{K_{22}}{2(K_{22} - a)} m_2.$$

Let $K_{11} = K_{22} = K_L$. Because of the structure of $\tilde{A}$, it can be verified that the rate of convergence for the estimation error (8) is $a - K_L$, which is the same as that of the Luenberger observer

[1] In general, $\bar{x}_2$ could be the convex combination of $\hat{x}_1$ and $\hat{x}_2$, i.e., $\bar{x}_2 = s_1\hat{x}_1 + s_2\hat{x}_2, s_1 + s_2 = 1, s_1, s_2 \in \mathbb{R}$.

(5). Moreover, assuming that constant noise m_1 and m_2 are equal, i.e., $m_1 = m_2 = m_0$, then $\bar{e}_2^\star = \frac{2K_L(K_L-a)-aK_{21}}{2(K_L-a)^2} m_0$. Interestingly, picking $K_{21} = \frac{2K_L(K_L-a)}{a}$, we obtain $\bar{e}_2^\star = 0$ for any unknown constant m_0, namely, the measurement noise can be completely rejected. When constant noise m_1 and m_2 are not equal, the choice $K_{21} = \frac{K_L(K_L-a)}{a}$ leads to $\bar{e}_2^\star = \frac{K_L}{2(K_L-a)} m_2$, which is a significant improvement (50%) over the case that agent 2 only has access to its own measurement (in which case $\bar{e}_2^\star = \frac{K_L}{K_L-a} m_2$). These properties cannot be achieved by using the Luenberger observer in (5).

For general measurement noises m_1 and m_2 (not necessarily constant), the H_∞ norm[2] from noise to the estimation error can be employed to study the noise effect. In fact, when $K_{21} \approx -4.75$, the H_∞ gain from noise m to the local estimate $\bar{e}_2$ achieves a minimum equal to 0.45, which is much smaller than that of the Luenberger observer in (5), which is 0.8, with equal rate of convergence ($K_L = 2, a = -0.5$).

It is important to point out that the observer proposed in this paper will also outperform the Luenberger observer in (5) when, in addition, agent 2 can transmit information to agent 1, i.e., the link between the two agents is bidirectional. Such an improvement is unique for the following two reasons. When the two agents are connected by a bidirectional link, our observer can be considered to be a bank of two observers providing a global estimate that averages the estimate of each individual observer. When the innovation terms "$-K_{21}(\hat{y}_1 - y_1)$" and "$-K_{12}(\hat{y}_2 - y_2)$" are missing, it can be shown that the effect of noise in the global estimate cannot be reduced – bank of observers currently available in the literature suffer from this shortcoming (see [25, Appendix D] for a proof of this claim). This suggests that the innovation terms in our interconnected observer are key. The second reason stems from the fact that our observer can be viewed as an "augmented-dimension observer" since, in general, it would have dimension Nn for a plant of dimension n. This property would contradict the well-known fact that an observer in the form (5) (or, in general, of the form (2)) minimizing the mean square estimation error under perturbations has necessarily the same dimension as the plant (see, e.g., [2, Section 4.2, Definition 4.3, and Theorem 4.5] and [25, Section IV.C]). However, when performance specifications (relative to the optimal observer) are added, which, in this paper, are formulated in terms of eigenvalue constraints, an n-dimensional observer may not be optimal. The augmented dimension (larger than the plant) is the key feature that enables our observer to outperform observers of the form (5), in particular, by mitigating the typical amplification of noise due to large gain required to speed up convergence.

As we show next, the idea behind the proposed interconnected observer illustrated in the example above generalizes to the case where N agents can measure the plant's output and share information over a graph.

III. Interconnected Observers

A. Notation and basic definitions

Given a matrix A with Jordan form $A = XJX^{-1}$, $\alpha(A) := \max\{Re(\lambda) : \lambda \in \text{eig}(A)\}$, where eig$(A)$ denotes the eigenspace of A; $\mu(A) := \max\{Re(\lambda)/2 : \lambda \in \text{eig}(A + A^\top)\}$; $|A| := \max\{|\lambda|^{\frac{1}{2}} : \lambda \in \text{eig}(A^\top A)\}$; $\kappa(A) := \min\{|X||X^{-1}| : A = XJX^{-1}\}$; A is dissipative if $A + A^\top < 0$. Given a vector $u \in \mathbb{R}^n$, $|u| := \sqrt{u^\top u}$. Given a function $m : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$, $|m|_\infty := \sup_{t \geq 0} |m(t)|$. The set of complex numbers is denoted by $\mathbb{C}$. The set of natural numbers is denoted by $\mathbb{N} := \{1, 2, 3, \cdots\}$. Given a symmetric matrix P, $\lambda_{\max}(P) := \max\{\lambda : \lambda \in \text{eig}(P)\}$ and $\lambda_{\min}(P) := \min\{\lambda : \lambda \in \text{eig}(P)\}$. For a continuous transfer function $\mathbb{C} \ni s \mapsto T(s) \in \mathbb{C}$, the H_∞ norm is defined as $||T||_\infty = \sup_{\omega \in \mathbb{R}} ||T(j\omega)||$, T is called stable if all its poles have negative real part, the dominant pole for a stable transfer function is the pole with largest real part, the rate of convergence of a closed-loop system with stable transfer function is defined by the absolute value of real part of the dominant pole. Given matrices A, B with proper dimensions, we define the operator $\text{He}(A, B) := A^\top B + B^\top A$; $A \otimes B$ defines the Kronecker product; and $A * B$ defines the Khatri-Rao product. Given a set S, the function card(S) defines the cardinality of the set S. A function $\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a class-$\mathcal{K}_\infty$ function, also written $\alpha \in \mathcal{K}_\infty$, if α is zero at zero, continuous, strictly increasing, and unbounded. A function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a class-$\mathcal{KL}$ function, also written $\beta \in \mathcal{KL}$, if it is nondecreasing in its first argument, nonincreasing in its second argument, $\lim_{r \to 0^+} \beta(r, s) = 0$ for each $s \in \mathbb{R}_{\geq 0}$, and $\lim_{s \to \infty} \beta(r, s) = 0$ for each $r \in \mathbb{R}_{\geq 0}$.

B. Preliminaries on graph theory

A directed graph (digraph) is defined as $\Gamma = (\mathcal{V}, \mathcal{E}, G)$. The set of nodes of the digraph are indexed by the elements of $\mathcal{V} = \{1, 2, \ldots, N\}$, and the edges are the pairs in the set $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$. Each edge directly links two nodes, i.e., an edge from i to j, denoted by (i, j), implies that agent i can send information to agent j. The adjacency matrix of the digraph Γ is denoted by $G = (g_{ij}) \in \mathbb{R}^{N \times N}$, where $g_{ij} = 1$ if $(i, j) \in \mathcal{E}$, and $g_{ij} = 0$ otherwise. A digraph is undirected if $g_{ij} = g_{ji}$ for all $i, j \in \mathcal{V}$. The in-degree and out-degree of agent i are defined by $d^{in}(i) = \sum_{j=1}^N g_{ji}$ and $d^{out}(i) = \sum_{j=1}^N g_{ij}$. The in-degree matrix D is the diagonal matrix with entries $D_{ii} = d^{in}(i)$, for all $i \in \mathcal{V}$. The set of indices corresponding to the neighbors that can send information to the i-th agent is denoted by $\mathcal{I}(i) := \{j \in \mathcal{V} : (j, i) \in \mathcal{E}\}$.

C. Observer structure and basic properties

The general form of the proposed observer consists of N *interconnected observers* with output given by the average over a graph of the states of the individual observers.[3] Specifically, consider a network of N agents defined by a graph $\Gamma = (\mathcal{V}, \mathcal{E}, G)$. For the estimation of the plant's state, a local state observer using information from its neighbors is attached to each agent. More precisely, for each $i \in \mathcal{V}$, the agent i runs a local state observer given by

$$\begin{aligned} \dot{\hat{x}}_i &= A\hat{x}_i - \sum_{j \in \mathcal{I}(i)} K_{ij}(\hat{y}_j - y_j), \\ \hat{y}_i &= C\hat{x}_i, \quad \bar{x}_i = \frac{1}{\text{card}(\mathcal{I}(i))} \sum_{j \in \mathcal{I}(i)} \hat{x}_j, \end{aligned} \tag{11}$$

where $\hat{x}_i$ denotes the state variable of the observer, $\bar{x}_i$ is the local estimate of the plant's state x, and y_i denotes the measurement of y in (1) taken by the i-th agent under measurement noise m_i, that is $y_i = Cx + m_i$. The information

[2] By "H_∞ norm" we mean the $\mathcal{L}_2$ gain from m to e, which is the induced 2-norm of the complex matrix transfer function from m to e.

[3] More general linear combinations defining $\bar{x}_i$ are possible, *i.e.*, $\bar{x}_i = \sum_{j \in \mathcal{I}(i)} \eta_j \hat{x}_j$ with $\eta_j \in \mathbb{R}$ for all j and $\sum_{j \in \mathcal{I}(i)} \eta_j = 1$.

that the i-th agent obtains from its neighbors are the values of $\hat{x}_j$'s and y_j's for each $j \in \mathcal{I}(i)$. The collection of local state observers in (11) connected via the graph Γ defines the proposed *interconnected observer*.

To analyze the properties of interconnected observers, define for each $i \in \mathcal{V}$, $e_i := \hat{x}_i - x$ and the associated vector $e := (e_1, \ldots, e_N)$. Furthermore, define the local estimation error $\bar{e}_i := \bar{x}_i - x$, the global estimation error vector $\bar{e} := (\bar{e}_1, \ldots, \bar{e}_N)$, and the noise vector $m := (m_1, \ldots, m_N)$. Then, it follows that

$$\begin{aligned} \dot{e}_i &= Ae_i - \sum_{j\in\mathcal{I}(i)} K_{ij}Ce_j + \sum_{j\in\mathcal{I}(i)} K_{ij}m_j, \\ \bar{e}_i &= \frac{1}{\mathrm{card}(\mathcal{I}(i))} \sum_{j\in\mathcal{I}(i)} e_j, \end{aligned} \tag{12}$$

which can be rewritten in the compact form

$$\begin{aligned} \dot{e} &= (I_N \otimes A - (\mathcal{K} * G^\top)(I_N \otimes C))e + (\mathcal{K} * G^\top)m, \\ \bar{e} &= (D^{-1} \otimes I_n)(G^\top \otimes I_n)e, \end{aligned} \tag{13}$$

where G is the adjacency matrix, D is the in-degree matrix,

$$\mathcal{K} = \begin{bmatrix} K_{11} & K_{12} & \cdots & K_{1N} \\ K_{21} & K_{22} & \cdots & K_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ K_{N1} & K_{N2} & \cdots & K_{NN} \end{bmatrix}, \tag{14}$$

and the Khatri-Rao product $\mathcal{K} * G^\top$ is such that $\mathcal{K}$ is treated as $N \times N$ block matrices with K_{ij}'s as blocks. Define

$$\begin{aligned} &\mathcal{A} := I_N \otimes A - (\mathcal{K} * G^\top)(I_N \otimes C), \\ &\mathcal{B} := \mathcal{K} * G^\top, \mathcal{C} := (D^{-1} \otimes I_n)(G^\top \otimes I_n). \end{aligned} \tag{15}$$

Then, the transfer function from measurement noise m to error $\bar{e}$ is given by $T(s) = \mathcal{C}(sI - \mathcal{A})^{-1}\mathcal{B}$. For the purpose of designing the proposed interconnected observer, each agent is self-connected, i.e., $(i, i) \in \mathcal{E}$. Therefore, we have $\mathrm{tr}(D) \geq N$.

Remark 3.1: The matrix $I_N \otimes A$ is a block diagonal matrix with matrix A in each of the N diagonal blocks (of dimension $n \times n$). The matrix $\mathcal{K} * G^\top$ defines the gain matrix for the graph, while $(D^{-1} \otimes I_n)(G^\top \otimes I_n)$ generates the estimation matrix for each agent by averaging the local estimates.

It can be verified that, under a detectability condition, interconnected observers can be designed so that the origin of the error system in (13) is (exponentially) stable.

Proposition 3.2: For the plant (1) *with measurement noise $m_i \equiv 0$ for each agent i, if the pair (A, C) is detectable, then, for any $N \in \mathbb{N}$, there exists a digraph Γ with adjacency matrix G and a gain $\mathcal{K}$ such that the matrix $\mathcal{A}$ is Hurwitz and the resulting system* (13) *has its origin exponentially stable.*

Proof: For any $N \in \mathbb{N}$, consider $G = I_N$. Then it follows that $\dot{e}_i = (A - K_{ii}C)e_i$ for each $i \in \mathcal{V}$. Under the assumption that the pair (A, C) is detectable, immediately we know that, for each $i \in \mathcal{V}$, there exists K_{ii} such that $A - K_{ii}C$ is Hurwitz. Therefore, the resulting $\mathcal{A}$ is Hurwitz. ■

D. $\mathcal{KL}$ characterization of performance and robustness

In this section, the performance and robustness properties of observers are characterized in terms of $\mathcal{KL}$ bounds. More precisely, given an observer with estimation error e, we are interested in bounds of the form

$$|e(t)| \leq \beta(|e(0)|, t) + \varphi(|m|_\infty) \qquad \forall t \geq 0,$$

where $t \mapsto e(t)$ is a solution to the error system, β is a class-$\mathcal{KL}$ function, and φ is a class-$\mathcal{K}_\infty$ function. To establish and compare this property with that of the interconnected observers, the next result characterizes such bounds for the proposed observer so that it can be designed to outperform those due to a Luenberger observers.

Proposition 3.3: For the plant (1)*, assume the pair (A, C) is detectable. Let $N \in \mathbb{N}$ and a digraph $\Gamma = (\mathcal{V}, \mathcal{E}, G)$ be given. If there exists a gain $\mathcal{K}$ such that at least one of the following conditions are satisfied:*

1) *The matrix $\mathcal{A}$ is Hurwitz with distinct eigenvalues;*
2) *The matrix $\mathcal{A}$ is dissipative, i.e., for some $\bar{\alpha} > 0$, $\mathcal{A}^\top + \mathcal{A} \leq -2\bar{\alpha}I$;*
3) *There exists $P = P^\top > 0$ such that $\mathrm{He}(\mathcal{A}, P) \leq -2\bar{\alpha}P$ for some $\bar{\alpha} > 0$;*

then, there exist a class-$\mathcal{KL}$ function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and a class-$\mathcal{K}$ function $\varphi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that the solution $\bar{e}$ of (13) *from any $e(0) \in \mathbb{R}^{nN}$ satisfies*

$$|\bar{e}(t)| \leq \beta(|e(0)|, t) + \varphi(|m|_\infty) \quad \forall t \in \mathbb{R}_{\geq 0}. \tag{16}$$

In particular, the functions β and φ can be chosen, for all $s, t \geq 0$, as follows: if 1) *holds, then, $\beta(s,t) = \kappa(\mathcal{A})|\mathcal{C}|\exp(\alpha(\mathcal{A})t)s$, $\varphi(s) = \kappa(\mathcal{A})\frac{|\mathcal{B}||\mathcal{C}|}{|\alpha(\mathcal{A})|}s$; if* 2) *holds, then, $\beta(s,t) = |\mathcal{C}|\exp(\mu(\mathcal{A})t)s$, $\varphi(s) = \frac{|\mathcal{B}||\mathcal{C}|}{|\mu(\mathcal{A})|}s$; if* 3) *holds, then, $\beta(s,t) = \sqrt{c_p}|\mathcal{C}|\exp(-\lambda t)s$, $\varphi(s) = c_p\frac{|\mathcal{B}||\mathcal{C}|}{|\lambda|}s$, with $\lambda = \frac{\bar{\alpha}\lambda_{\min}(P)}{\lambda_{\max}(P)}$ and $c_p = \frac{\lambda_{\max}(P)}{\lambda_{\min}(P)}$.*

Proof: The proof can be found in [25, Appendix A]. ■

Proposition 3.3 provides a way to design parameters for the interconnected observer as follows. Recall that $\tilde{A}_L$ and K_L are defined in (3).

Theorem 3.4: For the plant (1) *with the Luenberger observer* (2) *and the interconnected observers* (11)*, let $N \in \mathbb{N}$ and a digraph Γ be given. If K_L is such that at least one of the following conditions are satisfied:*

1) *$\tilde{A}_L$ is Hurwitz with distinct eigenvalues, and there exists $\mathcal{K}$ such that $\alpha(\mathcal{A}) < \alpha(\tilde{A}_L)$ and $\kappa(\mathcal{A})\frac{|\mathcal{B}||\mathcal{C}|}{|\alpha(\mathcal{A})|} < \kappa(\tilde{A}_L)\frac{|K_L|}{|\alpha(\tilde{A}_L)|}$;*
2) *$\tilde{A}_L$ is dissipative, and there exists $\mathcal{K}$ such that $\mu(\mathcal{A}) < \mu(\tilde{A}_L)$ (or $\alpha(\mathcal{A}) < \alpha(\tilde{A}_L)$, respectively – see below c)) and $\frac{|\mathcal{B}||\mathcal{C}|}{|\mu(\mathcal{A})|} < \frac{|K_L|}{|\mu(\tilde{A}_L)|}$;*
3) *$\tilde{A}_L$ satisfies $\mathrm{He}(\tilde{A}_L, P_L) \leq -2\overline{\alpha}_L P_L$ for some $\overline{\alpha}_L > 0$ and $P_L = P_L^\top > 0$, and there exists $\mathcal{K}$ such that*
   3.1) *item* 3) *of Proposition 3.3 holds with $\overline{\alpha} > 0, P = P^\top > 0$,*
   3.2) *$\lambda := \frac{\overline{\alpha}\lambda_{\min}(P)}{\lambda_{\max}(P)} < \frac{\overline{\alpha}_L\lambda_{\min}(P_L)}{\lambda_{\max}(P_L)} =: \lambda_L$ and $c_p\frac{|\mathcal{B}||\mathcal{C}|}{|\lambda|} < c_{pL}\frac{|K_L|}{|\lambda_L|}$, with $c_p = \frac{\lambda_{\max}(P)}{\lambda_{\min}(P)}$ and $c_{pL} = \frac{\lambda_{\max}(P_L)}{\lambda_{\min}(P_L)}$;*

then, there exist $\beta \in \mathcal{KL}$ and $\varphi \in \mathcal{K}_\infty$ such that the solution $\bar{e}$ of (13) *from any $e(0) \in \mathbb{R}^{nN}$ satisfies*

a) *$|\bar{e}(t)| \leq \beta(|e(0)|, t) + \varphi(|m|_\infty)$ for all $t \geq 0$;*
b) *Given nonzero $e(0)$ and $e_L(0)$, $\exists t^\star \geq 0$ such that $\beta(|e(0)|, t) < \beta_L(|e_L(0)|, t)$ for all $t > t^\star$;*
c) *$\varphi(s) < \varphi_L(s)$, for all $s \neq 0$, $s \in \mathbb{R}_{\geq 0}$.*

In particular, the functions $\beta \in \mathcal{KL}$ and $\varphi \in \mathcal{K}_\infty$ can be chosen accordingly as in Proposition 3.3 while $\beta_L \in \mathcal{KL}$ and $\varphi_L \in \mathcal{K}_\infty$ can be chosen, for all $s, t \geq 0$, as follows: if 1) *holds, then $\beta_L(s,t) = \kappa(\tilde{A}_L)\exp(\alpha(\tilde{A}_L)t)s$, $\varphi_L(s) = \kappa(\tilde{A}_L)\frac{|K_L|}{|\alpha(\tilde{A}_L)|}s$; if* 2) *holds, then $\beta_L(s,t) = \exp(\mu(\tilde{A}_L)t)s$ (or $\beta_L(s,t) = \kappa(\tilde{A}_L)\exp(\alpha(\tilde{A}_L)t)s$, respectively), $\varphi_L(s) =$*

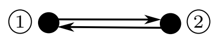
$\frac{|K_L|}{|\mu(\tilde{A}_L)|}s$; *if* 3) *holds, then* $\beta_L(s,t) = \sqrt{c_{pL}}\exp(-\lambda_L t)s$, $\varphi_L(s) = c_{pL}\frac{|K_L|}{|\lambda_L|}s$.

Proof: The proof follows from Proposition 3.3. Note that the Luenberger observer is a special case of the interconnected observer with $N = 1$. ■

Remark 3.5: Note that the boundedness property in item 2) in Theorem 3.4 guarantees that the rate of convergence of the interconnected observers is faster than or equal to that of a Luenberger observer by comparing the $\mathcal{KL}$ estimates they induce (which is a reasonable measure of performance when the $\mathcal{KL}$ functions are derived using similar bounding techniques).

The $\mathcal{KL}$ bounds established in Proposition 3.3 characterize a worse case property of the estimation error of the proposed observer, which can be compared to that of a Luenberger observer via Theorem 3.4. The following example illustrates this point.

Example 3.6: We revisit the motivational example in Section II and design an interconnected observer with $N = 2$ with an all-to-all graph as shown in Figure 2. Consider



$$\bar{x}_1 = \bar{x}_2 = \tfrac{1}{2}(\hat{x}_1 + \hat{x}_2)$$

Fig. 2. Two agents connected as a direct graph.

the case when two agents are experiencing common noises $m_1 = m_2 = m$. The transfer functions from m to e_L and from m to $\bar{e}$ (global) are given by[4] $T_L(s) = \frac{K_L}{s-a+K_L}$ and $T(s) = \mathcal{C}(sI - \mathcal{A})^{-1}\mathcal{B}$. In particular, the proposed observer takes the form

$$\begin{aligned} \dot{\hat{x}}_1 &= a\hat{x}_1 - K_{11}(\hat{y}_1 - y) - K_{12}(\hat{y}_2 - y), \\ \dot{\hat{x}}_2 &= a\hat{x}_2 - K_{22}(\hat{y}_2 - y) - K_{21}(\hat{y}_1 - y), \\ \hat{y}_1 &= \hat{x}_1, \quad \hat{y}_2 = \hat{x}_2, \quad \bar{x}_1 = \bar{x}_2 = \frac{\hat{x}_1 + \hat{x}_2}{2}. \end{aligned} \tag{17}$$

Then, we have the following result.

Proposition 3.7: Given $a, K_L \in \mathbb{R}$ *such that* $a \neq 0$ *and* $a - K_L < 0$, *then there exist* $K_{11}, K_{22}, K_{12}, K_{21} \in \mathbb{R}$ *such that the rate of convergence of the observer* (17) *is no smaller than that of the one induced by the Luenberger observer and the* H_∞ *norm of* T *is smaller than the* H_∞ *norm of* T_L, *i.e.,* $||T||_\infty < ||T_L||_\infty$.

Proof: The proof can be found in [25, Appendix B]. ■

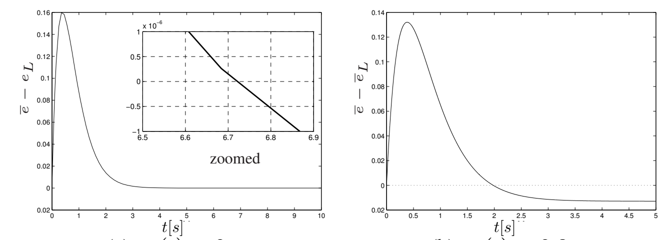
It should be noted that averaging the estimates of two uncoupled single Luenberger observers (one at each agent) does not lead to both faster convergence rate and smaller steady state error (see [25, Appendix D]). To perform a numerical comparison, we consider the case where $a = -0.5$ and $m : \mathbb{R}_{\geq 0} \to \mathbb{R}$ is a continuous bounded function. A Luenberger observer is designed following (5) to achieve a convergence rate of 2.5 and an H_∞ gain from m to e_L equal to 0.8, which leads to $K_L = 2$. For the interconnected observers (17), using Theorem 3.4, conditions 2) can be rewritten as

$$\begin{aligned} \alpha(\mathcal{A}) &\leq a - K_L, \\ \frac{\sqrt{2}}{2}\frac{\sqrt{(K_{11}+K_{12})^2 + (K_{22}+K_{21})^2}}{|\mu(\mathcal{A})|} &< \left|\frac{a}{a - K_L}\right|. \end{aligned} \tag{18}$$

[4] For the particular choice of parameters $K_{11} = K_{22} = K_L$ and $K_{12} = K_{21} = 0$, $||T||_\infty = ||T_L||_\infty$.

From solving (18), we pick parameters $K_{11} = 1.7896$, $K_{22} = 2.2278$, $K_{12} = 0.0538$, $K_{21} = -1.1633$. It can be verified that the eigenvalues of $\mathcal{A}$ according to this set of parameters are $-2.5087 \pm 0.1208i$. Moreover, $\mu(\mathcal{A}) = -1.9123$.

Now we perform simulations using these parameters and different measurement noises. With initial conditions $x(0) = 3$, $\overline{x}_1(0) = \overline{x}_2(0) = \overline{x}_L(0) = 5$, the first simulation is ran for measurement noise $m(t) \equiv 0$ and the resulting trajectories are shown in Figure 3(a). This figure shows that the interconnected observers converge at a faster rate compared to the Luenberger observer. In fact, item 2) of Theorem 3.4 holds with $t^\star \approx 6.7s$. Simulation results for $m(t) \equiv 0.3$ are shown in Figure 3(b). The behavior of the interconnected observers with constant noise is similar to that of with zero noise. It is worth to note that there is an improvement of the steady-state error by the interconnected observers since $\overline{e}^\star = 0.2272$, while the Luenberger observer gives $e_L^\star = 0.2400$. As shown in Figure 3(b), at around $t \approx 2s$, $\overline{e}$ becomes closer to 0 than $\overline{e}_L$ thereafter. To further explore the performance of the interconnected observers, we also consider measurement noise with different frequencies, *i.e.*, a low frequency noise $m(t) = 0.3 + 0.3\sin(20t)$ and a high frequency noise $m(t) = 0.3 + 0.3\sin(200t)$. The advantage of the interconnected observers lies on the properties of damped oscillatory behavior and smaller mean value of estimation error. Specifically, a numerical comparison of the estimation errors after transient is reported in the first two columns of Table I, which confirm the improvements guaranteed by the interconnected observers. △



(a) $m(t) \equiv 0$. (b) $m(t) \equiv 0.3$.

Fig. 3. Comparisons of estimation errors of the proposed observer and that of a Luenberger observer for different measurement noises with $N = 2$.

TABLE I
COMPARISON OF ESTIMATION ERROR ($\bar{e}$) OF THE OBSERVERS WITH MEASUREMENT NOISE OF DIFFERENT FREQUENCIES.

observer type	low freq. noise		high freq. noise		H_∞ from m to $\bar{e}$	
	mean $\bar{e}$	std $\bar{e}$	mean $\bar{e}$	std $\bar{e}$	Thm. 3.4	Thm. 3.10
Luenberger's	0.2419	0.0211	0.2395	0.0022	0.8000	0.8000
Interconnected	0.2286	0.0154	0.2268	0.0016	0.7600	0.5052
Improvement	**5.5%**	**27.0%**	**5.3%**	**27.3%**	**5.0%**	**38.1%**

E. Design via feasibility/optimization problems

The interconnected observers in (11) can be designed by solving feasibility and optimization problems that minimize the H_∞ gain of the transfer function from measurement noise m to estimation error $\bar{e}$ (global) or $\bar{e}_i$ (local) under the rate of convergence constraint. To formulate such problems, following [26], the error system in (13) is rewritten as

$$\dot{e} = A_e e + u, \quad y_e = C_e e + m, \quad z_\infty = \mathcal{X}e, \tag{19}$$

where $A_e = I_N \otimes A$, $C_e = -I_N \otimes C$, and the "input" u is assigned via $u = M_u y_e$ with $M_u = \mathcal{K} * G^\top$. Note that z_∞

denotes the overall estimation error (or the local estimation error) of the interconnected observers, *i.e.*, $z_\infty = \bar{e}$ with $\mathcal{X} = \mathcal{C}$ (or $z_\infty = \bar{e}_i$ with $\mathcal{X} = \mathcal{C}_i$). In the s-domain, the transfer function from m to z_∞ for (19) can be written as

$$T(s) = \mathcal{X}\big(sI - \mathcal{A}\big)^{-1}\mathcal{B} + \mathcal{D}, \tag{20}$$

where $\mathcal{A} = A_e + M_u C_e, \mathcal{B} = M_u$, and $\mathcal{D} = 0$. Within this setting, feasibility (*i.e.*, inequalities) and optimization problems associated with the design of the interconnected observers are formulated in the following sections.

1) Rate of convergence and H_∞ gain in terms of matrix inequalities: To guarantee that the rate of convergence of the interconnected observers is better (or no worse) than that of a Luenberger observer, the eigenvalues of the error system in (13) will be assigned to the left of the vertical line at $-\sigma$ in the s-plane, where σ is the rate of convergence for the Luenberger observer. Following [27], the error system (13) has all eigenvalues located to the left of $-\sigma$ on the s-plane if and only if there exists a matrix P_S such that

$$\text{He}(\mathcal{A}, P_S) + 2\sigma P_S < 0,\ \ P_S = P_S^\top > 0. \tag{21}$$

Note that (21) is nonlinear because of the cross term $P_S(\mathcal{K} * G^\top)$ obtained when expanding $P_S\mathcal{A}$. The following theorem provides an equivalent linear formulation and a sufficient condition for (21).

Proposition 3.8: Condition (21) *is satisfied if*

a) and only if $\text{He}(A_e, P_S) + C_e^\top M_p^\top + M_p C_e + 2\sigma P_S < 0$, $P_S = P_S^\top > 0$, *in which case* $M_u = P_S^{-1} M_p$;

b) the graph is all-to-all connected and there exists $h_1, h_2 \in \mathbb{R}$ *such that the following conditions hold:*

b.1) $h_1 + h_2 \geq \sigma$;

b.2) $P_i = P_i^\top > 0$ *for each* $i \in \mathcal{V}$

b.3) $\text{He}((A - K_{ii}C), P_i) + 2h_1 P_i < 0$ *for each* $i \in \mathcal{V}$;

b.4) $\begin{bmatrix} 2h_2P_1 & S_{12} & \cdots & S_{1N} \\ S_{12}^\top & 2h_2P_2 & \cdots & S_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ S_{1N}^\top & S_{2N}^\top & \cdots & 2h_2P_N \end{bmatrix} < 0$, *where* $S_{ij} = -(K_{ji}C)^\top P_j - P_i K_{ij} C$.

Proof: Letting $M_p = P_S M_u$, and using the definition of $\mathcal{A}$, inequality (21) can be written as

$$\text{He}(A_e, P_S) + C_e^\top M_p^\top + M_p C_e + 2\sigma P_S < 0,$$

with $P_S = P_S^\top > 0$. This proves item a). Now, assuming $b.1$)-$b.4$) with $h_1, h_2 \in \mathbb{R}$, note that the inequalities in $b.3$) can be rewritten as

$$\text{diag}(Q_1, \ldots, Q_N) + \text{diag}(2h_1P_1, \ldots, 2h_1P_N) < 0, \tag{22}$$

with $Q_i = \text{He}((A - K_{ii}C), P_i)$ for each $i \in \mathcal{V}$. By $b.2$), symmetry of the inequalities (22) and $b.4$), and the definition of negative symmetric matrices, the sum of the left terms of (22) and $b.4$) satisfies

$$\begin{bmatrix} Q_1 & S_{12} & \cdots & S_{1N} \\ S_{12}^\top & Q_2 & \cdots & S_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ S_{1N}^\top & S_{2N}^\top & \cdots & Q_N \end{bmatrix} + \begin{bmatrix} 2\bar{h}P_1 & 0 & \cdots & 0 \\ 0 & 2\bar{h}P_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 2\bar{h}P_N \end{bmatrix} < 0, \tag{23}$$

with $\bar{h} = h_1 + h_2$. Since $h_1 + h_2 \geq \sigma$, (21) is satisfied with $P_D = \text{diag}(P_1, \ldots, P_N)$. ■

Proposition 3.9: Conditions $b.1$)-$b.4$) *in Proposition 3.8 hold if and only if there exist* $h_1,\ h_2, \in \mathbb{R},\ Y_i,\ W_{ij},\ P_i$ *for* $i, j \in \mathcal{V}$ *and* $j \neq i$ *such that:*

a) $h_1 + h_2 \geq \sigma$,

b) $P_i = P_i^\top > 0$, *for each* $i \in \mathcal{V}$,

c) $\text{He}(A, P_i) - C^\top Y_i^\top - Y_i C + 2h_1 P_i < 0$, *for each* $i \in \mathcal{V}$,

d) $\begin{bmatrix} 2h_2P_1 & R_{12} & \cdots & R_{1N} \\ R_{21} & 2h_2P_2 & \cdots & R_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ R_{N1} & R_{N2} & \cdots & 2h_2P_N \end{bmatrix} < 0,$

where $R_{ij} = -C^\top W_{ji}^\top - W_{ij} C$.

The conditions $b.3$)-$b.4$) *in Proposition 3.8 hold with* $K_{ii} = P_i^{-1} Y_i$ *and* $K_{ij} = P_i^{-1} W_{ij}$ *for* $i, j \in \mathcal{V}$, $j \neq i$.

Proof: Let $Y_i = P_i K_{ii}$ and $W_{ij} = P_i K_{ij}$ for $i, j \in \mathcal{V}$ and $j \neq i$, then, $b.3$)-$b.4$) in Proposition 3.8 can be rewritten as

$$\text{He}(A, P_i) - C^\top Y_i^\top - Y_i C + 2h_1 P_i < 0$$

for each $i \in \mathcal{V}$ and

$$\begin{bmatrix} 2h_2P_1 & R_{12} & \cdots & R_{1N} \\ R_{21} & 2h_2P_2 & \cdots & R_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ R_{N1} & R_{N2} & \cdots & 2h_2P_N \end{bmatrix} < 0,$$

respectively. Therefore, c) and d) of Proposition 3.9 hold. ■

2) Minimization of H_∞ norm under rate of convergence constraint with fixed connectivity graph: We consider the design of interconnected observer over a fixed digraph $\Gamma = (\mathcal{V}, \mathcal{E}, G)$. The design specifications of our interest are the rate of convergence and the H_∞ gain from noise m to estimation errors $\bar{e}$ or e_i.

Theorem 3.10: Given a plant as in (1) *and a digraph* Γ, *the rate of convergence is larger than or equal to* σ *and the global H_∞ gain (respectively, the local H_∞ gain) from m to estimation error $\bar{e}$ in* (13) *(respectively, $\bar{e}_i$ in* (12)*) is minimized if and only if there exist matrices $\mathcal{K}$, P_S, and P_H such that the following optimization problem has a solution:*

$$\inf \gamma \tag{24a}$$

$$\textit{s.t.}\ \text{He}(\mathcal{A}, P_S) + 2\sigma P_S < 0, \tag{24b}$$

$$\begin{bmatrix} \text{He}(\mathcal{A}, P_H) & P_H\mathcal{B} & \mathcal{X}^\top \\ \mathcal{B}^\top P_H & -\gamma I & 0 \\ \mathcal{X} & 0 & -\gamma I \end{bmatrix} < 0, \tag{24c}$$

$$P_S = P_S^\top > 0,\ P_H = P_H^\top > 0, \tag{24d}$$

where $\mathcal{X} = \mathcal{C}$ (respectively, $\mathcal{X} = \mathcal{C}_i$ and $\mathcal{C}_i$ is the sub-matrix of $\mathcal{C}$ from the $(in - n + 1)$-th row to the (in)-th row).

Proof: From [28, Theorem 2.41], the H_∞ gain for a system from input to output with realization $T_1(s) = C_1(sI - A_1)^{-1}B_1$ is less than or equal to γ if and only if there exists some $P_H = P_H^\top > 0$ such that

$$\begin{bmatrix} \text{He}(A_1, P_H) & P_H B_1 & C_1^\top \\ B_1^\top P_H & -\gamma I & 0 \\ C_1 & 0 & -\gamma I \end{bmatrix} < 0, \tag{25}$$

Then, for system (13) with $T(s) = \mathcal{C}(sI - \mathcal{A})^{-1}\mathcal{B}$, we have that the global H_∞ gain from m to $\bar{e}$ is less than or equal to γ if and only if (25) holds with $A_1 = \mathcal{A}$, $B_1 = \mathcal{B}$ and $C_1 = \mathcal{C}$, which leads to (24c) with $\mathcal{X} = \mathcal{C}$. The same argument applies for $T_i(s) = \mathcal{C}_i(sI - \mathcal{A})^{-1}\mathcal{B}$ which leads to (24c) with $\mathcal{X} = \mathcal{C}_i$. Then, the proof finishes by adding constraint (21). ■

Remark 3.11: For a fixed connectivity graph, the optimization problem in (24) can be solved offline. Moreover, due to the form of the observer at each node as in (11), the information needed by each agent is what the neighbors provide through the connectivity graph. Therefore, the resulting observers for each agent are decentralized.

Note that the optimization problem (24) is not jointly convex over the variables (P_S, P_H, M_u). Moreover, it is nonlinear because of the existence of cross terms $P_H M_u$ and $P_S M_u$. In order to remove the nonlinearities and make the two constraints jointly convex, following [26], we reformulate the problem by seeking common solutions of P_S and P_H, and changing variables to $M_p := PM_u$. Using item $a)$ of Proposition 3.8 to rewrite the terms $\mathrm{He}(\mathcal{A}, P_S)$ and $\mathrm{He}(\mathcal{A}, P_H)$ in (24), we have the following result.

Theorem 3.12: Given a plant as in (1) *and a graph* Γ, *the rate of convergence is larger than or equal to* σ *and the global* H_∞ *gain (respectively, the local* H_∞ *gain) from* m *to estimation error* $\bar{e}$ *in* (13) *(respectively,* $\bar{e}_i$ *in* (12)*) is minimized if there exist* M_p *and* P *such that the following optimization problem (LMI) is feasible:*

$$
\begin{aligned}
&\inf \gamma \\
&\text{s.t.: } \mathrm{He}(A_e, P) + C_e^\top M_p^\top + M_p C_e + 2\sigma P < 0, \\
&\begin{bmatrix} \mathrm{He}(A_e, P) + C_e^\top M_p + M_p^\top C_e & M_p & \mathcal{X}^\top \\ M_p^\top & -\gamma I & 0 \\ \mathcal{X} & 0 & -\gamma I \end{bmatrix} < 0, \\
&P = P^\top > 0,
\end{aligned}
$$

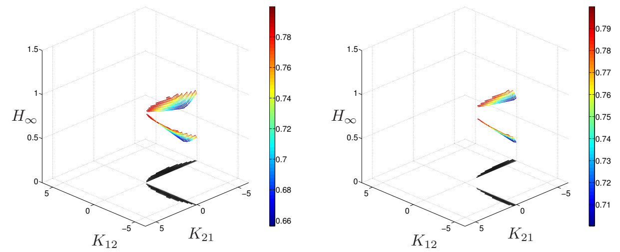
where $\mathcal{X} = \mathcal{C}$ *(respectively,* $\mathcal{X} = \mathcal{C}_i$ *and* $\mathcal{C}_i$ *is the sub-matrix of* $\mathcal{C}$ *from the* $(in - n + 1)$*-th row to the* (in)*-th row).*

Remark 3.13: The resulting observer gain matrix from Theorem 3.12 is given by $M_u = P^{-1} M_p$. By making the optimization problem linear and convex, a global optimizer is guaranteed. However, asking for common solution of $P_H = P_D$ may eliminate a better feasible solution to the original optimization problem in (24).

Following [29], it is possible to formulate an equivalent convex optimization problem to the one in Theorem 3.12 but with noncommon P_D and P_H matrices, see [25, Appendix F].

Example 3.14: We revisit the motivational example with connectivity graph as in Figure 1. To further indicate the improvement obtained by the proposed observer, we choose $K_{11} = K_{22} = K_L = 2$, and $K_{21} = -0.5K_L = -1$. The resulting local H_∞ gain from m to $\bar{e}_2$ is 0.55, which is smaller than that of the Luenberger observer, which is 0.8. If instead the connectivity graph in Figure 2 is considered, we can further optimize the parameters by solving the optimization problem (24). Feasible parameters for (24) are found using the solver PENBMI [30]. For $K_{11} \approx -6.7215$, $K_{22} \approx 10.7215$, $K_{12} \approx -13.2202$, $K_{21} \approx 5.7537$, the resulting H_∞ gain is ≈ 0.5051, which is $\approx 36.86\%$ smaller than that of Luenberger observer (which is 0.8 with $K_L = 2$). This improvement and the improvement obtained when using Theorem 3.4 are listed in the last two columns of Table I.

In fact, when the rate of convergence specification is $\sigma = 2.5$, and the H_∞ gain from m to $\bar{e}$ is restricted to be less than or equal to 0.8, then, by letting $K_{11} = 2$ and $K_{22} = 2$, we can find the feasible region for K_{12} and K_{21} as shown in Figure 4(a). Moreover, if the rate of convergence is required to be $\sigma = 3.0$ with the same H_∞ constraint, then, by letting $K_{11} = 2.5$ and $K_{22} = 2.5$, we obtain the feasible region for K_{12} and K_{21} as shown in Figure 4(b). As the figure suggests, faster rate of convergence leads to a smaller feasible region for the observer parameters. More importantly, for a single Luenberger observer, there is no feasible solution for rate of convergence larger than or equal to 3.0 and global H_∞ gain less than 0.8.



(a) Regions for rate of convergence equal 2.5 ($K_{12} = K_{21} = 2$).

(b) Regions for rate of convergence equal 3.0 ($K_{12} = K_{21} = 2.5$).

Fig. 4. Feasible regions for observer parameters subject to different rate of convergence specification and global H_∞ gain less than 0.8.

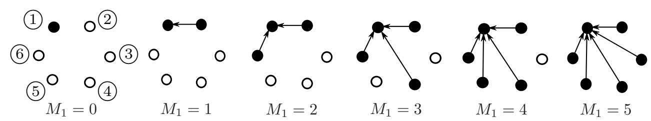


Fig. 5. Different graph structures for agent 1 with $N = 6$.

TABLE II
COMPARISON OF LOCAL H_∞ NORMS FROM NOISE m TO $\bar{e}_1$ WITH DIFFERENT NUMBER OF INCOMING EDGES FOR AGENT 1.

	number of non-self edges (M_1)					
	0	1	2	3	4	5
Local H_∞	0.80	0.45	0.34	0.28	0.25	0.22
Improvement	**0.00%**	**43.8%**	**57.5%**	**65.0%**	**68.8%**	**72.5%**

Now, for the same plant, consider digraphs with 6 agents where the edges are defined as in Figure 5. In all cases, each agent is self connected. Let M_1 denote the number of non-self edges for agent 1, e.g., when $M_1 = 0$ as shown in Figure 5, it is implied that $G = I_6$, while when $M_1 = 5$, $G = \begin{bmatrix} g_1 & g_2 \end{bmatrix}$, $g_1 = [1\ 1_5^\top]^\top$ and $g_2 = [0\ I_5]^\top$. Let the rate of convergence specification be $\sigma = 2.5$. Then, the local H_∞ norms from noise $m = (m_1, \ldots, m_6)$ to estimation error $\bar{e}_1$ at agent 1 for the cases in Figure 5 are shown in Table II. From case $M_1 = 0$ to case $M_1 = 1$, the improvement is significant; in fact, when an incoming edge is added to agent 1, the local H_∞ is improved by 43.8% when compared to the case where a single Luenberger observer is used at agent 1. When two agents provide information to agent 1 ($M_1 = 2$), the improvement is approximately 57.5%, while when three and four agents communicate to agent 1, the improvement grows to approximately 65% and 69% ($M_1 = 4$), respectively. $\triangle$

Example 3.15 (second order plant): First, we consider a second-order plant given as in (1) with $A = \begin{bmatrix} -5/2 & 1/10 \\ 4/100 & -3 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 2 \end{bmatrix}$. For a given Luenberger observer with $K_L = [1.5 \quad -0.16]^\top$, its rate of convergence is -3.34 and its H_∞ norm from measurement noise m to estimation error e_L is approximately equal 0.34. With the interconnected observers for $N = 2$ connected via an all-to-all connectivity graph, we obtain that the optimal global H_∞ norm from measurement noise m to estimation error

$\bar{e}$ is approximately 0.05 and the optimal local H_∞ norm from m to $\bar{e}_1$ (or $\bar{e}_2$) is 0.03 with $M_u = [v_1\ v_2]$, where $v_1 = [10.3834\ \ -1.6019\ \ -10.7581\ 1.5963]^\top$ and $v_2 = [7.1992\ \ -1.2410\ \ -7.3028\ 1.2426]^\top$. The resulting global and local H_∞ gains are $\approx 87.88\%$ and $\approx 91.43\%$ smaller than that of Luenberger observers, respectively.

Then, we consider a second-order plant with oscillatory behavior given as in (1) with $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 0 \end{bmatrix}$. For a given Luenberger observer with $K_L = [2\ \ 0]^\top$, its rate of convergence is -1 and its H_∞ norm from measurement noise m to estimation error e_L is equal 2. With the interconnected observers with $N = 2$ connected via an all-to-all connectivity graph, by formulating the problem according to (19), the optimization problem in Theorem 3.10 is solved and the gain matrix is found as $M_u = [v_1\ v_2]$, where $v_1 = [7.9503\ \ -9.9554\ \ -5.9424\ 9.0014]^\top$ and $v_2 = [-5.9324\ 9.1143\ 7.9605\ -9.8426]^\top$. Its corresponding global H_∞ norm from m to $\bar{e}$ is ≈ 1.4125 and its local H_∞ norm from m to $\bar{e}_1$ (or $\bar{e}_2$) is ≈ 1. Comparing to the Luenberger observer, the global H_∞ norm is decreased by $\approx 29.4\%$ and the local H_∞ norm is decreased by $\approx 50.0\%$. $\triangle$

The improvements on the local H_∞ gain guaranteed by the proposed interconnected observers in the examples above are justified by the fact that the sufficient condition given in the upcoming Section III-F are satisfied; see Theorem 3.20 and below it, where these examples are revisited.

3) Minimization of H_∞ norm under rate of convergence constraint with optimized connectivity graph: For interconnected observers whose digraph has not yet been specified, a natural question to ask is whether there exists a digraph that minimizes the number of links between agents for the given specifications. In applications, such minimizations could potentially lower the cost of a distributed system as it could reduce the number of agents and communication links. The following result provides a sufficient and necessary condition for such optimization problem.

Theorem 3.16: For the error system (13)*, the rate of convergence is larger than or equal to σ and the global H_∞ norm (respectively, the local H_∞ norm) from noise m to estimation error $\bar{e}$ in* (13) *(respectively, $\bar{e}_i$ in* (12)*) is less than or equal to $\gamma^\star$ over a digraph Γ with minimized number of edges if and only if there exist matrices $\mathcal{K}$, G, P_S, and P_H such that the following optimization problem has a solution:*

$$\inf \operatorname{tr}(D) \tag{26a}$$

$$s.t.\ \operatorname{He}(\mathcal{A}, P_S) + 2\sigma P_S < 0, \tag{26b}$$

$$\begin{bmatrix} \operatorname{He}(\mathcal{A}, P_H) & P_H\mathcal{B} & \mathcal{X}^\top \\ \mathcal{B}^\top P_H & -\gamma^\star I & 0 \\ \mathcal{X} & 0 & -\gamma^\star I \end{bmatrix} < 0, \tag{26c}$$

$$P_S = P_S^\top > 0,\ P_H = P_H^\top > 0, \tag{26d}$$

where $\mathcal{X} = \mathcal{C}$ (respectively, $\mathcal{X} = \mathcal{C}_i$).

Proof: Following the proof of Theorem 3.10, the global H_∞ gain over a digraph Γ is less than or equal to $\gamma^\star$ if and only if (25) holds with $A_1 = \mathcal{A}$, $B_1 = \mathcal{B}$, $C_1 = \mathcal{C}$, $\gamma = \gamma^\star$ and $P_H = P_H^\top > 0$. The same argument applies to the local H_∞ gain. Moreover, the rate of convergence condition is satisfied if and only if (26b) holds. Since $\operatorname{tr}(D) = \sum_{i=1}^N \sum_{j=1}^N g_{ij}$, where $g_{ij} = 1$ indicates there is an edge from node j to node i, then the number of edges of the graph is minimized if and only if $\operatorname{tr}(D)$ is minimized. ■

The constraints in (26b) and (26c) are nonlinear and not jointly convex. By changing variables, the nonlinear constraints in (26b) and (26c) can be linearized as a function of Q and P.

Theorem 3.17: For the error system (13)*, the rate of convergence is larger than or equal to σ and the global H_∞ norm (respectively, the local H_∞ norm) from noise m to estimation error $\bar{e}$ in* (13) *(respectively, $\bar{e}_i$ in* (12)*) is less than or equal to $\gamma^\star$ over a digraph Γ with minimized number of communication links if there exist matrices $\mathcal{K}$, G, and P such that the following optimization problem is feasible:*

$$\inf \operatorname{tr}(D) \tag{27a}$$

$$s.t.\ \operatorname{He}(I_N \otimes A, P) + Z + 2\sigma P < 0, \tag{27b}$$

$$\begin{bmatrix} \operatorname{He}(I_N \otimes A, P) + Z & Q & \mathcal{X}^\top \\ Q^\top & -\gamma^\star I & 0 \\ \mathcal{X} & 0 & -\gamma^\star I \end{bmatrix} < 0, \tag{27c}$$

$$P = P^\top > 0, \tag{27d}$$

where $Q = P(\mathcal{K} * G^\top)$, $Z = -Q(I_N \otimes C) - (I_N \otimes C)^\top Q^\top$, and $\mathcal{X} = \mathcal{C}$ (respectively, $\mathcal{X} = \mathcal{C}_i$).

Proof: Let $\mathcal{K}$, G and P be solutions of the optimization problem (27). Since the matrix $\mathcal{K} * G^\top$ is such that $Q = P(\mathcal{K} * G^\top)$, using $P = P^\top$ and the definition of $\mathcal{A}$ in (15), we have

$$\begin{aligned} &\operatorname{He}(I_N \otimes A, P) - Q(I_N \otimes C) - (I_N \otimes C)^\top Q^\top \\ &= (I_N \otimes A)^\top P + P(I_N \otimes A) \\ &\quad - (I_N \otimes C)^\top (\mathcal{K} * G^\top)^\top P^\top - P(\mathcal{K} * G^\top)(I_N \otimes C) \\ &= \operatorname{He}(\mathcal{A}, P). \end{aligned}$$

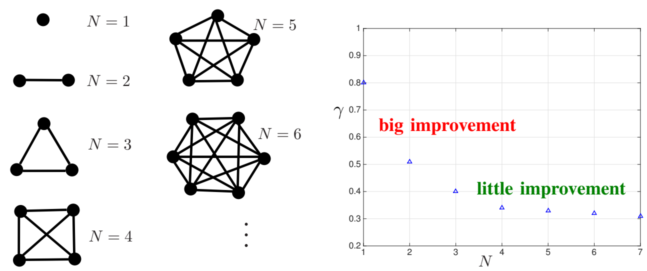
Then, $\mathcal{K}$, G, $P_S = P$ and $P_H = P$ satisfy (26). ■

Remark 3.18: The results above define the graph via the resulting G. The resulting $\mathcal{K}$ and G from (27) satisfies $\mathcal{K} * G^\top = P^{-1}Q$, which may not be unique.

Example 3.19: Consider the scalar plant in (4) with $a = -0.5$ as in Example 3.14, which can represent the dynamics of a mobile agent whose state is to be estimated using multiple sensors either fixed or mobile (in relative coordinates). Suppose that the rate of convergence specification is $\sigma = 2.5$. When using the graph that is all-to-all as shown in Figure 6(a), it is natural to ask the effect that the number of agents has on the improvement of the global H_∞ norm. As shown in Figure 6(b), the resulting global H_∞ gain is reduced as the number of agents N grows. These results are obtained following Theorem 3.16. The improvement is summarized in Table III. Note that the improvement is less significant for $N > 6$. In particular, the table indicates that if the global H_∞ gain is required to be less than or equal to 0.40, then, as shown in Figure 6(b),

TABLE III
COMPARISON OF GLOBAL H_∞ NORMS FROM NOISE m TO $\bar{e}$ WITH DIFFERENT NUMBER OF AGENTS UNDER ALL-TO-ALL CONNECTION.

	number of agents (N)						
	1	2	3	4	5	6	7
global H_∞	0.80	0.51	0.40	0.34	0.33	0.32	0.31
improvem't	**0.00%**	**36.3%**	**50.0%**	**57.5%**	**58.8%**	**60.0%**	**61.3%**
local H_∞	0.80	0.38	0.24	0.23	0.21	0.20	0.19
improvem't	**0.00%**	**52.5%**	**70.0%**	**71.3%**	**73.8%**	**75.0%**	**76.3%**



(a) Graph structures with all to all connections (self connection links are not shown).

(b) The global H_∞ norm from noise m to estimation error $\bar{e}$ with respect to the number of agents.

Fig. 6. The influence of the number of agents on the H_∞ gain from noise m to estimation error $\bar{e}$.

the least number of agents needed is three[5]. For the same scalar plant with three interconnected observers, according to Theorem 3.16, we establish a relationship between tr(D) and the global H_∞ gain from m to estimation error $\bar{e}$ in Table IV. In particular, for tr(D) smaller than six, there is no improvement in the H_∞ gain when compared to that of Luenberger observers. Moreover, the table indicates that, with three interconnected observers, if the global H_∞ gain is required to be less than or equal to 0.6, then the minimum number of links required in the connectivity graph Γ is seven. △

TABLE IV
COMPARISON OF GLOBAL H_∞ NORMS FROM NOISE m TO $\bar{e}$ WITH DIFFERENT CONNECTIVITY GRAPH WITH $N=3$.

	tr(D)			
	6	7	8	9
global H_∞	0.64	0.53	0.43	0.40
improvement	**20.0%**	**33.8%**	**46.3%**	**50.0%**

F. A sufficient condition guaranteeing smaller local H_∞ gain

In this section, we are interested in conditions on the plant (1) for which it is possible to design interconnected observers that, for a given rate of convergence σ, have local H_∞ gains smaller than when a single Luenberger observer is used at each agent. Note that the local H_∞ gain affects the quality of the estimates obtained at each node. These estimates can be computed efficiently and in a decentralized manner using local information, while computing the global estimate requires additional algorithms – see [25, Section IV.B]. The following result provides one such condition.

Theorem 3.20: Given $\sigma \geq 0$, suppose K_L is such that the eigenvalues of the error system (3) *of the Luenberger observer* (2) *for the plant* (1) *are located in the region $\mathcal{D} = \{s \in \mathcal{C}_0 : Re(s) < -\sigma\}$, and the H_∞ gain from m to e_L is $\gamma_L > 0$. If there exist $\tilde{\alpha} \in \mathbb{R}$ and $P = P^\top > 0$ such that*

$$\begin{bmatrix} \mathrm{He}(A - K_L C, P) & P K_L C & -\tilde{\alpha} I_n \\ C^\top K_L^\top P & -I_n & (1+\tilde{\alpha}) I_n \\ -\tilde{\alpha} I_n & (1+\tilde{\alpha}) I_n & -I_n \end{bmatrix} < 0, \quad (28)$$

then, for every $N \in \mathbb{N},\ N > 1$, there exist a digraph Γ and a gain $\mathcal{K}$ for N interconnected observers in (11) *such that the error system* (13) *has its eigenvalues in $\mathcal{D}$ and the local H_∞ gain from m to associated $\bar{e}_i$ for all agents are less than or equal to γ_L. Moreover, for at least $N-1$ agents, the local H_∞ gain from m to associated $\bar{e}_i$ is strictly less than γ_L.*

[5]The optimization problem related to the examples shown in this paper are solved by PENBMI [30].

Proof: For any $N > 1$, let the digraph Γ have adjacency matrix

$$G_N = \begin{bmatrix} 1 & 1_{N-1}^\top \\ 0 & I_{N-1} \end{bmatrix}. \quad (29)$$

This choice of G indicates that agent 1 can share information with all other agents. Moreover, for each $i \in \mathcal{V}$, let T_i be the transfer function from m to $\bar{e}_i$. Take $N = 2$ and $K_{11} = K_{22} = K_L$, $K_{12} = 0$, and K_{21} to be determined later. Then, the interconnected observers in (11) reduce to

$$\begin{aligned} \dot{\hat{x}}_1 &= A\hat{x}_1 - K_L(\hat{y}_1 - y_1), \\ \dot{\hat{x}}_2 &= A\hat{x}_2 - K_L(\hat{y}_2 - y_2) - K_{21}(\hat{y}_1 - y_1), \\ \hat{y}_1 &= C\hat{x}_1,\ \hat{y}_2 = C\hat{x}_2,\quad \bar{x}_1 = \hat{x}_1,\ \bar{x}_2 = \frac{\hat{x}_1 + \hat{x}_2}{2}, \end{aligned} \quad (30)$$

with associated error system as in (13) with

$$\mathcal{A} = \begin{bmatrix} A - K_L C & 0 \\ -K_{21} C & A - K_L C \end{bmatrix},\ \mathcal{B} = \begin{bmatrix} K_L & 0 \\ K_{21} & K_L \end{bmatrix}.$$

If K_L is such that (2) has its eigenvalues in $\mathcal{D} = \{s \in \mathcal{C}_0 : Re(s) < -\sigma\}$, then, due to the block matrix form of $\mathcal{A}$, the eigenvalues of $\mathcal{A}$ are also in $\mathcal{D}$. Now, suppose (28) holds with $\alpha \in \mathbb{R}$ and $P = P^\top > 0$. Then, if (28) is treated as an H_∞ constraint, equivalently, we have

$$\left|\left| -\tilde{\alpha}(sI - \tilde{A}_L)^{-1} K_L C + (1+\tilde{\alpha}) I \right|\right|_\infty < 1. \quad (31)$$

Therefore, the transfer function $T_2(s) = \mathcal{C}_2(sI - \mathcal{A})^{-1}\mathcal{B}$ satisfies

$$T_2 = \frac{1}{2}[I\ \ I] \begin{bmatrix} sI - \tilde{A}_L & 0 \\ K_{21} C & sI - \tilde{A}_L \end{bmatrix}^{-1} \begin{bmatrix} K_L & 0 \\ K_{21} & K_L \end{bmatrix}.$$

By using the inversion identity for a block matrix (inversion lemma), it follows that

$$\begin{bmatrix} sI - \tilde{A}_L & 0 \\ K_{21} C & sI - \tilde{A}_L \end{bmatrix}^{-1} = \begin{bmatrix} (sI - \tilde{A}_L)^{-1} & 0 \\ F & (sI - \tilde{A}_L)^{-1} \end{bmatrix},$$

where $F = -(sI - \tilde{A}_L)^{-1} K_{21} C (sI - \tilde{A}_L)^{-1}$ Then, by assigning $K_{21} = \tilde{\alpha} K_L$, T_2 can be simplified as

$$T_2 = \left[\ \tfrac{1}{2} T_L - \tfrac{1}{2}\tilde{\alpha} T_L C T_L + \tfrac{1}{2}\tilde{\alpha} T_L \quad \tfrac{1}{2} T_L\ \right],$$

where $T_L(s) = (sI - \tilde{A}_L)^{-1} K_L$. Therefore, we obtain

$$||T_2||_\infty \leq \frac{1}{2}||(1+\tilde{\alpha})T_L - \tilde{\alpha} T_L C T_L||_\infty + \frac{1}{2}||T_L||_\infty.$$

Using (31), it follows that $||T_2||_\infty < ||T_L||_\infty = \gamma_L$. Now consider for any $N > 1, N \in \mathbb{N}$, with digraph whose adjacency matrix is G_N, it follows that the transfer function T_i from noise m to $\bar{e}_i$ satisfies $T_i = T_2$ for all $i \in \mathcal{V}, i \neq 1$. Therefore, $||T_i||_\infty < \gamma_L$ for all $i \in \mathcal{V}, i \neq 1$. ■

Note that condition (28) is a property on the plant for a given K_L; basically, an H_∞ inequality as in (24c). Next, we illustrate this condition in the examples throughout the paper.

Example 3.21: For the scalar plant (4) with the Luenberger observer (5), the transfer function in the s-domain from m to e_L is given by $T_L(s) = \frac{K_L}{s - a + K_L}$. Since (28) is an LMI with respect to P and $\tilde{\alpha}$, its feasibility can be easily verified, *e.g.*, for $a = -0.5$ and $K_L = 2$, $P = 0.47$ and $\tilde{\alpha} = -0.5$ solve (28). Therefore, for the plant (4), there exist interconnected observers such that at least $N-1$ local H_∞ gains are smaller than $\gamma_L = 0.8$ with $K_L = 2$. This justifies the improvement shown in the motivational example as in Table I. △

Example 3.22: We revisit the systems in Example 3.15. For the first system discussed in Example 3.15, the improvement is justified by the fact that condition (28) in Theorem 3.20 holds with $\tilde{\alpha} = -0.3241$ and $P = 0.1I$. △

While it may be possible to get further improvement by designing the gains of the interconnected observers as in the design of Kalman filters, it should be noted that the tradeoff between performance and robustness affects general Kalman filters; see [25, Section IV.C] for a discussion on this.

IV. Conclusion

In contrast to standard observers for linear time-invariant systems, interconnected observers have the capability of attaining fast rate of convergence rate without necessarily jeopardizing robustness to measurement noise in the H_∞ sense. The comparison between $\mathcal{KL}$ bounds between interconnected and Luenberger observers leads to checkable conditions that can be used for design – though potentially conservative. When solved for specific systems, the stated feasibility and optimization problems lead to significant improvements, when compared to single Luenberger observers. Such improvement is guaranteed by the satisfaction of an LMI condition. While the optimization of the number of internal observers and the connectivity graph are not necessarily linear and convex, numerical results for a particular plant indicate that the improvement obtained in robustness is significant only up to a finite number of such internal observers.

References

[1] D. Luenberger, "An introduction to observers," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 596–602, December 1971.

[2] H. Kwakernaak and R. Silvan, *Linear Optimal Control Systems*. John WIley & Sons, Inc., 1972.

[3] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in Systems and Control Theory*. Siam, 1994.

[4] H. Li and M. Fu, "A linear matrix inequality approach to robust H-infinity filtering," *IEEE Transactions on Signal Processing*, vol. 45, no. 9, pp. 2338–2350, 1997.

[5] H. J. Marquez, "A frequency domain approach to state estimation," *Journal of the Franklin Institute*, vol. 340, no. 2, pp. 147–157, 2003.

[6] P. Ioannou and J. Sun, *Robust adaptive control*. Prentice Hall, 1996.

[7] A. Astolfi and L. Praly, "Global complete observability and output-to-state stability imply the existence of a globally convergent observer," *Mathematics of Control, Signals, and Systems (MCSS)*, vol. 18, no. 1, pp. 32–65, 2006.

[8] V. Andrieu, L. Praly, and A. Astolfi, "High gain observers with updated gain and homogeneous correction terms," *Automatica*, vol. 45, no. 2, pp. 422 – 428, 2009.

[9] J. H. Ahrens and H. K. Khalil, "High-gain observers in the presence of measurement noise: A switched-gain approach," *Automatica*, vol. 45, no. 4, pp. 936–943, 2009.

[10] H. Lei, J. Wei, and W. Lin, "A global observer for observable autonomous systems with bounded solution trajectories," in *Proc. of 44th IEEE Conference on Decision and Control*, Dec. 2005, pp. 1911–1916.

[11] A. A. Ball and H. K. Khalil, "Analysis of a nonlinear high-gain observer in the presence of measurement noise," in *Proc. of American Control Conference*, July 2011, pp. 2584–2589.

[12] R. G. Sanfelice and L. Praly, "On the performance of high-gain observers with gain adaptation under measurement noise," *Automatica*, vol. 47, no. 10, pp. 2165–2176, 2011.

[13] H. Shim, A. Tanwani, and Z. Ping, "Back-and-forth operation of state observers and norm estimation of estimation error," in *Proc. of 51st IEEE Conference on Decision and Control*, Dec 2012, pp. 3221–3226.

[14] J. Cortes, "Distributed Kriged Kalman filter for spatial estimation," *IEEE Transactions on Automatic Control*, vol. 54, no. 12, pp. 2816–2827, 2009.

[15] R. Olfati-Saber, "Distributed Kalman filter with embedded consensus filters," in *Proc. of 44th IEEE Conference on Decision and Control, European Control Conference.*, 2005, pp. 8179–8184.

[16] D. P. Spanos, R. Olfati-Saber, and R. M. Murray, "Approximate distributed Kalman filtering in sensor networks with quantifiable performance," in *Proc. of 4th International Symposium on Information Processing in Sensor Networks*, 2005.

[17] P. Alriksson and A. Rantzer, "Distributed Kalman filtering using weighted averaging," in *Proc. of 17th International Symposium on Mathematical Theory of Networks and Systems*, 2006.

[18] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proc. of 46th IEEE Conference on Decision and Control*, 2007, pp. 5492–5498.

[19] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri, "Distributed Kalman filtering based on consensus strategies," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 4, pp. 622–633, 2008.

[20] Y. Zhao, Z. Duan, G. Wen, and G. Chen, "Distributed H_∞ consensus of multi-agent systems: a performance region-based approach," *International Journal of Control*, vol. 85, no. 3, pp. 332–341, 2012.

[21] S. S. Kia, J. Cortes, and S. Martinez, "Dynamical average consensus under limited control authority and privacy requirements," *International Journal of Robust and Nonlinear Control*, 2013.

[22] Y. Hong, G. Chen, and L. Bushnell, "Distributed observers design for leader-following control of multi-agent networks," *Automatica*, vol. 44, no. 3, pp. 846–850, 2008.

[23] J. Hu and X. Hu, "Nonlinear filtering in target tracking using cooperative mobile sensors," *Automatica*, vol. 46, no. 12, pp. 2041 – 2046, 2010.

[24] J. Hu, X. Hu, and T. Shen, "Cooperative shift estimation of target trajectory using clustered sensors," *Journal of Systems Science and Complexity*, vol. 27, no. 3, pp. 413–429, 2014.

[25] Y. Li and R. G. Sanfelice, "Interconnected observers for robust decentralized estimation with performance guarantees and optimized connectivity graph," University of California, Santa Cruz, Technical Report, 2015, http://arxiv.org/abs/1503.08706.

[26] C. Scherer, P. Gahinet, and M. Chilali, "Multiobjective output-feedback control via LMI optimization," *IEEE Transactions on Automatic Control*, vol. 42, no. 7, pp. 896–911, July 1997.

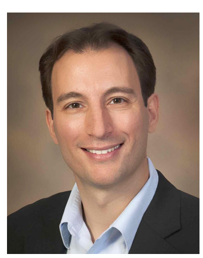
[27] M. Chilali and P. Gahinet, "H_∞ design with pole placement constraints: an LMI approach," *IEEE Transactions on Automatic Control*, vol. 41, no. 3, pp. 358–367, March 1996.

[28] C. Scherer, "The riccati inequality and state-space H_∞-optimal control," Ph.D. dissertation, Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades der Bayerischen Julius Maximilians-Universitat Wurzburg, 1990.

[29] Y. Ebihara and T. Hagiwara, "A dilated LMI approach to robust performance analysis of linear time-invariant uncertain systems," *Automatica*, vol. 41, no. 11, pp. 1933–1941, 2005.

[30] M. Kocvara and M. Stingl, "PENNON: A code for convex nonlinear and semidefinite programming," *Optimization Methods and Software*, vol. 18, no. 3, pp. 317–333, 2003.

Yuchun Li received his B.S. and M.S. degree in Mechanical Engineering from Zhejiang University, Hangzhou, China, in 2007 and 2010, respectively. Currently, he is pursuing a Ph.D. in the Hybrid Systems Laboratory in the Department of Computer Engineering at the University of California, Santa Cruz. His research interests include modeling, stability, observer design, control and robustness analysis of hybrid systems.



Ricardo G. Sanfelice is an Associate Professor of Computer Engineering, University of California at Santa Cruz, CA, USA. He received his M.S. and Ph.D. degrees in 2004 and 2007, respectively, from the University of California, Santa Barbara. During 2007 and 2008, he was a Postdoctoral Associate at the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology and visited the Centre Automatique et Systemes at the Ecole de Mines de Paris for four months. Prof. Sanfelice is the recipient of the 2013 SIAM Control and Systems Theory Prize, the National Science Foundation CAREER award, the Air Force Young Investigator Research Award, and the 2010 IEEE Control Systems Magazine Outstanding Paper Award. His research interests are in modeling, stability, robust control, observer design, and simulation of nonlinear and hybrid systems with applications to power systems, aerospace, and biology.