

Multiplier-based Observer Design for Large-Scale Lipschitz Systems

Ming Jin, Han Feng, and Javad Lavaei

Abstract—Observer design for nonlinear systems with incomplete state observations is of practical significance. Despite numerous existing works and recent developments in handling multiple types of nonlinearities, it is still an open challenge to reduce conservatism of synthesis conditions for systems with large Lipschitz constants, and to improve computational efficiency for complex real-world dynamics. To this end, this study presents a multiplier-based approach that is capable of determining an asymptotically stable observer for a large class of highly nonlinear and large-scale systems. These key advantages are due to an informative quadratic constraint on the nonlinear dynamics. Both the present and the state-of-the-art methods are evaluated in a benchmark example and a case study on the dynamic power system state estimation, where the proposed approach exhibits an imperative trade-off between non-conservatism and computational tractability, establishing its viability for real-world large-scale nonlinear systems.

I. INTRODUCTION

Real-world physical systems, such as power grids and autonomous vehicles, are often large-scale and nonlinear [1]–[4]. Because the entire state of a real-world system may not be accessible due to economical/technological constraints, a fundamental problem in system analysis and control is how to estimate the state from measurements, which is known as observer design [1]. This study focuses on the important case¹

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{f}(\mathbf{x}) + \mathbf{B}\mathbf{u} \\ \mathbf{y} = \mathbf{C}\mathbf{x} \end{cases}, \quad (1)$$

where \mathbf{A} , \mathbf{B} and \mathbf{C} are matrices of appropriate dimensions, $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$ are the state and the observation vectors, and $\mathbf{u} \in \mathbb{R}^n$ is a known control input (this implies that the observer has access to the output of the controller). We further assume that the nonlinear part $\mathbf{f}(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ satisfies (local) Lipschitz continuity, namely for all \mathbf{x}_1 and \mathbf{x}_2 in the subset $\mathcal{B} \in \mathbb{R}^n$,

$$\|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_2)\| \leq \gamma \|\mathbf{x}_1 - \mathbf{x}_2\|, \quad (2)$$

where γ is known as the Lipschitz constant, and $\mathbf{f}(\cdot)$ may also capture uncertain parameters in the presence of measurement errors and changing environments. While observer design can be easily integrated in the controller design for linear time-invariant systems, it presents unique challenges for nonlinear

systems and has been the subject of investigation in numerous works [2], [5]–[14]. Despite recent developments in handling multiple types of nonlinearities [2], [10], [11], [15], the estimation problem is still open and several difficulties remain, namely: (1) to reduce conservativeness of the design methods (e.g., for highly nonlinear systems with larger Lipschitz constants); and (2) to improve computational scalability to analyze physical systems at larger scales.

Motivated by these pressing needs, this study proposes a multiplier-based approach for the observer design problem, borrowing powerful ideas from the framework of integral quadratic constraints (IQC) [16] and dissipation theory [17]. The IQC framework is celebrated for its capability to encompass a variety of real-life uncertainties and for computational efficiency in analyzing large-scale systems. The main obstacle in applying IQC, nonetheless, is the lack of available mechanisms to characterize multiple-input multiple-output (MIMO) Lipschitz nonlinearities. Existing tools are either potentially very conservative (e.g., sector-bound IQC) or impose stringent conditions (e.g., Zames-Falb IQC [18] and its MIMO extension require the nonlinear function to be the derivative of a convex function [19]). To address this issue, we propose a new quadratic constraint on Lipschitz nonlinearities by exploiting inherent structures of MIMO Lipschitz functions. With this informative constraint in place, in tandem with the introduction of multipliers, this study enables the design of observers for a major class of large-scale and highly nonlinear systems. The multiplier-based framework can also incorporate uncertainties/nonlinearities that have practical implications, such as uncertain and time-varying parameters.

The paper is organized as follows. An overview of observer design methods for nonlinear systems is provided in Sec. II. The multiplier-based design method is introduced in Sec. III. To numerically evaluate the developed method and compare it with the state-of-the-art approaches, experiments on a benchmark example and a power grid state estimation problem are conducted in Sec. IV. Concluding remarks are given in Sec. V.

II. RELATED WORK

Observer design for nonlinear systems has been widely investigated in the literature [1], [2], [5], [8]–[13], [15]. Several types of observers have been studied, including Luenberger-like observers with a single gain matrix [6], [8], [14] or multiple gain matrices [2], [10], [20]. In many cases, the problem can be reduced to solving linear matrix inequalities (LMIs) by applying the Young’s inequality [8], [13], Riccati equations [7], [13], or via introducing multipliers

[†]This work was supported by the ONR grants N00014-17-1-2933 and N00014-15-1-2835, DARPA grant D16AP00002, and AFOSR grant FA9550-17-1-0163.

*M. Jin, H. Feng and J. Lavaei are with the Department of Industrial Engineering and Operations Research, University of California, Berkeley. J. Lavaei is also with the Tsinghua-Berkeley Shenzhen Institute. Emails: {jinning, han.feng, lavaei}@berkeley.edu

¹This formulation is general in the sense that one can always linearize the system dynamics around an equilibrium point.

based on the S-procedure [2], [3], [9], [10], [21], [22]. Asymptotic observers for nonlinear systems with globally Lipschitz nonlinearities have been developed in [7], [8], [20], [23]. The technique in [20] has also been extended to systems with single-input single-output slope-restricted nonlinearities [20], MIMO nonlinearities satisfying a monotonicity condition [10], and time-varying nonlinearities satisfying incremental quadratic constraints [2]. These methods all employ the observer structure that has multiple gain matrices to add more degrees of freedom. A general description of a nonlinear observer with an “output injection” form was first given in [24], and further analyzed by [25] in an incremental stability framework. Despite novel attempts to exploit the decoupled nature of Lipschitz nonlinearities by introducing decomposition matrices [10], [13], existing LMI synthesis conditions remain restrictive for many nonlinear systems.

To overcome this difficulty, a new linear parameter-varying (LPV) formulation was introduced in [14], which was shown to substantially reduce the conservatism of prior methods [3], [9], [13], [22]. However, as remarked in that paper, a major bottleneck of the proposed LPV method is its high computational complexity (e.g., 2^{n^2} LMIs need to be solved simultaneously for an n -dimensional nonlinear vector), which significantly limits its application to large-scale systems. This hurdle is removed in the present study, based on a reformulation of the Lipschitz condition as a quadratic constraint that depends on auxiliary variables. By introducing multipliers using the S-procedure [21], this constraint can be incorporated into a single LMI, which only scales quadratically in terms of the dimension of the nonlinearities. This leads to a less restrictive observer design for large-scale real-world physical systems, as demonstrated in numerical experiments.

III. MULTIPLIER-BASED OBSERVER DESIGN

The starting point of this analysis is a less conservative constraint on general vector-valued smooth functions.

A. Characterization of Lipschitz functions

Lipschitz continuity (2) implies uniform continuity. Unless otherwise stated, we use the Euclidean norm in our analysis. The property (2) can be expressed as a point-wise quadratic constraint for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{B}$:

$$\begin{bmatrix} \mathbf{x}_1 - \mathbf{x}_2 \\ \mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_2) \end{bmatrix}^\top \begin{bmatrix} \gamma^2 \mathbf{I}_n & \\ & -\mathbf{I}_m \end{bmatrix} [\star] \geq 0, \quad (3)$$

where we use \star to denote the symmetric component. The above constraint, nevertheless, can be sometimes very conservative, because it does not explore the structure of the problem [14]. To understand this, consider the function

$$\mathbf{f}(x_1, x_2) = \left[\frac{1}{1+e^{-0.5x_1}} - ax_1, \sin(x_2) \right]^\top, \quad (4)$$

where $x_1, x_2 \in \mathbb{R}$ and $|a| \leq 0.1$ is a deterministic but unknown parameter with a bounded magnitude. Clearly, to satisfy (2), we need to specify that $\gamma \geq 1$ (i.e., the function has Lipschitz constant 1). However, this characterization is too

general in this case, because it ignores the *non-homogeneity* of f_1 and f_2 , as well as the *sparsity* of the inputs x_1 and x_2 . Indeed, f_1 only depends on x_1 with its slope restricted to $[-0.1, 0.6]$ for all possible values $|a| \leq 0.1$, and f_2 only depends on x_2 with its slope restricted to $[-1, 1]$. In the context of controller design, the non-homogeneity of control outputs often arises from physical constraints and domain specifications, and the sparsity of inputs is inherent to decentralized/distributed control. To explicitly address these requirements, we state the following quadratic constraint, which was partially inspired by [14].

Lemma 1 (Quadratic constraint for Lipschitz functions). *For a vector-valued function $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ that is differentiable with bounded partial derivatives on \mathcal{B} (i.e., $\underline{b}_{ij} \leq \partial_j f_i(\mathbf{x}) \leq \bar{b}_{ij}$, for all $\mathbf{x} \in \mathcal{B}$), the following quadratic constraint is satisfied for all $\lambda_{ij} \geq 0$, $i \in [m]$, $j \in [n]$, and $\mathbf{x}, \mathbf{z} \in \mathcal{B}$:²*

$$\begin{bmatrix} \mathbf{x} - \mathbf{z} \\ \mathbf{q} \end{bmatrix}^\top \underbrace{\begin{bmatrix} \text{diag}(\{\sum_i \lambda_{ij} (\bar{c}_{ij}^2 - c_{ij}^2)\}) & \Sigma(\{\lambda_{ij}, c_{ij}\})^\top \\ \Sigma(\{\lambda_{ij}, c_{ij}\}) & \text{diag}(\{-\lambda_{ij}\}) \end{bmatrix}}_{\mathbf{M}(\boldsymbol{\lambda}, \boldsymbol{\gamma})} [\star] \geq 0, \quad (5)$$

where $\mathbf{q} = [q_{11}, \dots, q_{1n}, \dots, q_{m1}, \dots, q_{mn}]^\top \in \mathbb{R}^{mn}$ is a function of \mathbf{x} and \mathbf{z} , $\{\lambda_{ij} c_{ij}\}$ follows the same index order as \mathbf{q} , $\Sigma(\{\lambda_{ij}, c_{ij}\})^\top = [\text{diag}(\{\lambda_{1j} c_{1j}\}) \ \cdots \ \text{diag}(\{\lambda_{mj} c_{mj}\})] \in \mathbb{R}^{n \times mn}$, $c_{ij} = \frac{1}{2}(\underline{b}_{ij} + \bar{b}_{ij})$, $\bar{c}_{ij} = \bar{b}_{ij} - c_{ij}$, and \mathbf{q} is related to the output of \mathbf{f} by the constraint:

$$\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{z}) = \underbrace{[\mathbf{I}_m \otimes \mathbf{1}_{1 \times n}]}_{\mathbf{w}} \mathbf{q}, \quad (6)$$

where \mathbf{I}_m is the $m \times m$ identity matrix, $\mathbf{1}_{1 \times n}$ is the $1 \times n$ all-one matrix, \otimes denotes the Kronecker product, and $\boldsymbol{\gamma}$ in $\mathbf{M}(\boldsymbol{\lambda}, \boldsymbol{\gamma})$ denotes the set of bounds \underline{b}_{ij} and \bar{b}_{ij} on partial derivatives.

Proof: See Appendix A. \square

This bound is a direct consequence of standard tools in real analysis. To understand this result, it can be verified that (2) is equivalent to:

$$\sum_{ij} \lambda_{ij} \left((\bar{c}_{ij}^2 - c_{ij}^2)(x_j - y_j)^2 + 2c_{ij}q_{ij}(x_j - z_j) - q_{ij}^2 \right) \geq 0, \quad (7)$$

for all $\lambda_{ij} \geq 0$, with $f_i(\mathbf{x}) - f_i(\mathbf{z}) = \sum_{j=1}^n q_{ij}$. Since (7) holds for all $\lambda_{ij} \geq 0$, it is equivalent to the condition that $(\bar{c}_{ij}^2 - c_{ij}^2)(x_j - z_j)^2 + 2c_{ij}q_{ij}(x_j - z_j) - q_{ij}^2 \geq 0$ for all $i \in [m]$, $j \in [n]$, which is a direct result of the bounds associated with the partial derivatives of f_i . To illustrate the application of this idea, we apply it to (4), where $\underline{b}_{11} = -0.1$, $\bar{b}_{11} = 0.6$, $\underline{b}_{22} = -1$, $\bar{b}_{22} = 1$ while the other bounds ($\underline{b}_{12}, \bar{b}_{12}, \underline{b}_{21}, \bar{b}_{21}$) are zero. This clearly yields a more informative constraint than simply relying on the Lipschitz constraint (3). In fact, for a differentiable Lipschitz function, we have $\bar{b}_{ij} = -\underline{b}_{ij} =$

²We use $[n]$ to denote the set $\{1, \dots, n\}$.

γ , and by limiting the choice of $\lambda_{ij} = \begin{cases} \lambda & \text{if } i = 1 \\ 0 & \text{if } i \neq 1 \end{cases}$, (7)

is reduced to (3). However, as illustrated in this example, the quadratic constraint in Lemma 1 can incorporate richer information about *input sparsity* and *output structures*, and thus it can often reduce conservatism in practice.

To simplify the mathematical treatment, we have focused on differentiable functions in Lemma 1; nevertheless, the analysis can be extended to non-differentiable but Lipschitz continuous functions using the notions of generalized gradient [26, Chap. 2]. In brief, by re-assigning the bounds on partial derivatives to uniform bounds on the set of generalized partial derivatives, the constraint (5) can be directly applied.

This constraint is in the form similar to an IQC. In relation to existing IQCs, it has wider applications to characterize smooth functions. The Zames-Falb IQC introduced in [18] has been widely used for single-input single-output (SISO) function $f : \mathbb{R} \rightarrow \mathbb{R}$, but it requires the function to be monotone with its slope restricted to $[\alpha, \beta]$ with $\alpha \geq 0$ (i.e., $0 \leq \alpha \leq \frac{f(x)-f(y)}{x-y} \leq \beta$ for all $x \neq y$). Its MIMO extension holds true only if $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is restricted to be the *gradient of a convex real-valued function* [19]. As for the sector IQC, the vector version is in fact (3). By contrast, the quadratic constraint in Lemma 1 can be applied to non-monotone, vector-valued Lipschitz functions, where existing IQCs cannot be used.

B. Observer design

With the newly developed quadratic constraint on smooth functions in place, this subsection illustrates the design of an observer for the system (1). We focus on the classical Luenberger observer [8], [13], [14], [27]:

$$\dot{\hat{\mathbf{x}}} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{f}(\hat{\mathbf{x}}) + \mathbf{L}(\mathbf{y} - \mathbf{C}\hat{\mathbf{x}}) + \mathbf{B}\mathbf{u}, \quad (8)$$

where $\hat{\mathbf{x}} \in \mathbb{R}^n$ is the estimated state and $\mathbf{L} \in \mathbb{R}^{n \times m}$ is to be designed. The dynamics of the estimation error $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$ can be described by:

$$\dot{\mathbf{e}} = (\mathbf{A} - \mathbf{L}\mathbf{C})\mathbf{e} + \mathbf{f}(\mathbf{x}) - \mathbf{f}(\hat{\mathbf{x}}). \quad (9)$$

The objective is to design \mathbf{L} such that the estimation error $\mathbf{e}(t)$ asymptotically decreases to $\mathbf{0}$. The robust version of the error dynamics (9) can be written as

$$\begin{cases} \dot{\mathbf{e}} = (\mathbf{A} - \mathbf{L}\mathbf{C})\mathbf{e} + \phi \\ \phi = \mathbf{w} + \mathbf{v} \\ \mathbf{w} = \mathbf{f}(\mathbf{x}) - \mathbf{f}(\hat{\mathbf{x}}) \end{cases} \quad (10)$$

where $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}} \in \mathbb{R}^n$ is the estimation error, $\mathbf{v} \in \mathbb{R}^n$ is the perturbation noise with bounded \mathcal{L}_2 energy ($\int_0^T \|\mathbf{v}(t)\|_2^2 dt < \infty$ for all $T > 0$), and $\mathbf{w} \in \mathbb{R}^n$ arises from the system nonlinearities. We introduce an internal variable $\mathbf{q} \in \mathbb{R}^{n^2}$ that is related to \mathbf{w} by the underdetermined equality $\mathbf{w} = \mathbf{W}\mathbf{q}$, where $\mathbf{W} = [\mathbf{I}_n \otimes \mathbf{1}_{1 \times n}] \in \mathbb{R}^{n \times n^2}$ is defined in Lemma 1. Because $\mathbf{f}(\cdot)$ is Lipschitz continuous, the signals \mathbf{e} and \mathbf{q} jointly satisfy the quadratic constraint (5). The objective is to search for \mathbf{L} such that the observer is asymptotically stable.

This property can be derived from input-output stability at all time $T \geq 0$, i.e.,

$$\int_0^T \|\mathbf{e}(t)\|_2^2 dt \leq \rho^2 \int_0^T \|\mathbf{v}(t)\|_2^2 dt + c, \quad (11)$$

where c is a finite constant and ρ is a finite upper bound constant for the \mathcal{L}_2 gain. In particular, as long as \mathbf{v} has finite energy, the right-hand side is a bounded constant for all T . The asymptotic stability follows by allowing T to approach infinity. To this end, define the feasibility problem $\text{SDP}(\mathbf{P}, \mathbf{L}, \boldsymbol{\lambda}, \rho, \gamma)$ as follows:

$$\text{SDP}(\mathbf{P}, \mathbf{L}, \boldsymbol{\lambda}, \rho, \gamma) : \begin{bmatrix} \mathcal{O}(\mathbf{P}, \mathbf{L}, \boldsymbol{\lambda}, \rho, \gamma) & \mathcal{S}(\mathbf{P}) \\ \mathcal{S}(\mathbf{P})^\top & -\rho \mathbf{I}_n \end{bmatrix} \preceq 0, \quad (12)$$

where $\mathbf{P} = \mathbf{P}^\top \succ 0$ is positive definite, γ incorporates the upper/lower bounds on partial derivatives (e.g., Lipschitz upper bound of $\mathbf{f}(\cdot)$), and $\mathcal{O}(\mathbf{P}, \mathbf{L}, \boldsymbol{\lambda}, \rho, \gamma)$ is given by

$$\begin{bmatrix} (\mathbf{A} - \mathbf{L}\mathbf{C})^\top \mathbf{P} + \mathbf{P}(\mathbf{A} - \mathbf{L}\mathbf{C}) + \frac{1}{\rho} \mathbf{I}_n & \mathbf{P}\mathbf{W} \\ \mathbf{W}^\top \mathbf{P} & \mathbf{0}_{n^2, n^2} \end{bmatrix} + \mathbf{M}(\boldsymbol{\lambda}, \gamma),$$

and

$$\mathcal{S}(\mathbf{P}) = \begin{bmatrix} \mathbf{P} \\ \mathbf{0}_{n^2, n} \end{bmatrix},$$

where $\mathbf{M}(\boldsymbol{\lambda}, \gamma)$ is defined in (5) with non-negative multipliers $\boldsymbol{\lambda} = \{\lambda_{ij}\}$ for $i \in [n]$ and $j \in [n]$. We will show next that the above condition can be used to synthesize an asymptotically stable observer.

Theorem 2. *Let $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a bounded nonlinear function of the state. Assume that $\mathbf{f}(\cdot)$ is γ -Lipschitz with bounded partial derivatives on \mathcal{B} (i.e., $\underline{b}_{ij} \leq \partial_j f_i(\mathbf{x}) \leq \bar{b}_{ij}$ and $|\underline{b}_{ij}|, |\bar{b}_{ij}| \leq \gamma$ for all $\mathbf{x} \in \mathcal{B}$, $i \in [n]$ and $j \in [n]$). If there exist a scalar $\rho > 0$, $\mathbf{L} \in \mathbb{R}^{n \times m}$ and $\mathbf{P} = \mathbf{P}^\top \succ 0$ such that $\text{SDP}(\mathbf{P}, \mathbf{L}, \boldsymbol{\lambda}, \rho, \gamma)$ is feasible, then the observer parameterized by \mathbf{L} is stable (i.e., it satisfies (11)).* \square

The proof follows from a standard dissipation argument and is relegated to the appendix. Because $\text{SDP}(\mathbf{P}, \mathbf{L}, \boldsymbol{\lambda}, \rho, \gamma)$ is a bilinear matrix inequality with the term $\mathbf{P}\mathbf{L}$, we can replace it with a new variable $\mathbf{R} \in \mathbb{R}^{n \times m}$. Then, the observer gain \mathbf{L} is given by $\mathbf{P}^{-1}\mathbf{R}$. This makes $\text{SDP}(\mathbf{P}, \mathbf{R}, \boldsymbol{\lambda}, \rho, \gamma)$ quasiconvex, in the sense that it reduces to a standard LMI with fixed ρ . To solve the resulting problem numerically, we start with a small \mathcal{L}_2 gain ρ and gradually increase it until a solution $(\mathbf{P}, \mathbf{R}, \boldsymbol{\lambda})$ is found. Each iteration (i.e., LMI for a given set of ρ) can be solved efficiently by interior-point methods, which have been implemented in SDPT3, Mosek, and SeDuMi. As an alternative to performing a search on ρ , more sophisticated methods for solving the generalized eigenvalue optimization problem can be employed [21].

IV. EXPERIMENTS

In this section, we evaluate the developed methodology on a benchmark system, in addition to a real-world example of power grid dynamic state estimation.

A. Flexible link robot

This example considers a one-link manipulator with revolute joints actuated by a DC motor [7], [13], [14]. The system dynamics can be modeled as (1) with the parameters

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -48.6 & -1.25 & 48.6 & 0 \\ 0 & 0 & 0 & 1 \\ 19.5 & 0 & -19.5 & 0 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0 \\ 21.6 \\ 0 \\ 6 \end{bmatrix},$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \mathbf{f}(\mathbf{x}) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -3.33 \sin(x_3) \end{bmatrix},$$

where the states are the motor position and velocity as well as the link position and velocity. The Lipschitz constant γ for this system is 3.33. By solving LMI (12) with fixed $\gamma = 3.33$ (Lipschitz constant) and $\rho = 1.5$ (certified \mathcal{L}_2 gain), the stabilizing gain matrix is given by

$$\mathbf{L}^\top = \begin{bmatrix} 3.37 & -47.25 & 1.05 & 21.75 \\ 0 & 56.62 & 124.54 & 304.20 \end{bmatrix}. \quad (13)$$

A simulation with an open-loop excitation $u(t) = \sin(2\pi t)$ is shown in Fig. 1 to illustrate the observer's tracking ability. Furthermore, to demonstrate the superiority of the proposed approach in dealing with highly nonlinear systems, we increase the Lipschitz constant of the system (i.e., increase the mass of the links) until the feasibility LMI condition cannot be solved, and use it to evaluate the existing methods [9], [13], [14]. The results are listed in Table I (see [14] for a discussion of the potential bias in comparing LMI (50) and LMI (28) in [13]). Compared with [13], our method and [9], [14] can deal with systems with very large Lipschitz constants. Even though the methods are comparable in this example, [9] does not apply to MIMO nonlinearities, and [14] needs to solve an exponential number of LMIs, which is not scalable to larger systems; on the contrary, our method can be applied for MIMO nonlinearities with large Lipschitz constants while maintaining computational tractability.

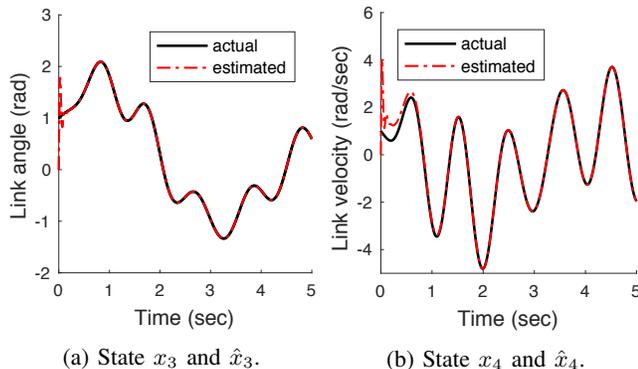


Fig. 1: Tracking performance of the observer with the gain matrix (13) and the initial condition set to be 0.

TABLE I: Comparison results for different LMI methods for the flexible link robot example.

	LMI (50) in [13]	LMI (28) in [13]	LMI (9) in [9]	LMI (35) in [14]	LMI (12)
γ_{\max}	0.99	48.5	$\approx 10^7$	$\approx 10^{10}$	$\approx 10^7$

* Strict positive/negative definite inequalities imposed with 10^{-6} margin.

B. Power grid dynamic state estimation

In this subsection, we study the power grid phase and frequency dynamic estimation. The New England Power System (NE-PS) under analysis is shown in Fig. 2. Under the prescribed information structure, the observer can only access phases and frequencies at generators G1, G2 and G3, and the objective is to determine the states at the remaining geographically separated counterparts.

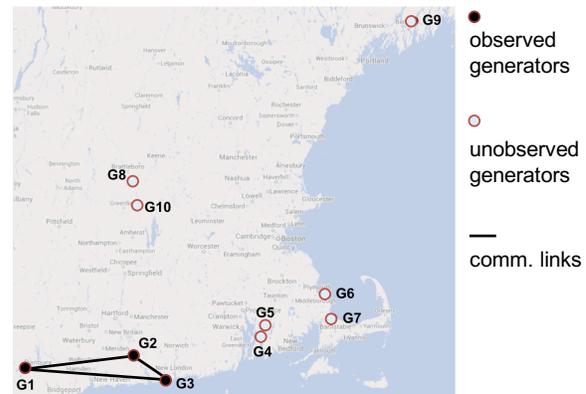


Fig. 2: Information structure in NE-PS under study, where only states at G1, G2 and G3 can be accessed.

By defining θ_i as the voltage angle at a generator bus i (in rad), the physics of power systems are modeled by the per-unit swing equation:

$$Q_i \ddot{\theta}_i + K_i \dot{\theta}_i = P_{M_i} - P_{E_i}, \quad (14)$$

where P_{M_i} is the mechanical power input to the generator at bus i (in p.u.), P_{E_i} is the electrical active power injection at bus i (in p.u.), Q_i is the inertia coefficient of the generator at bus i (in p.u.-sec²/rad), and K_i is the damping coefficient of the generator at bus i (in p.u.-sec/rad). The electrical real power injection P_{E_i} depends on the voltage angle difference in a nonlinear way, as governed by the AC power flow equation:

$$P_{E_i} = \sum_{j=1}^n |V_i| |V_j| (G_{ij} \cos(\theta_i - \theta_j) + S_{ij} \sin(\theta_i - \theta_j)),$$

where n is the number of generators in the system, G_{ij} and S_{ij} are the conductance and susceptance of the transmission line that connects buses i and j , V_i is the voltage phasor at bus i , and $|V_i|$ is its voltage magnitude. Because the conductance G_{ij} is typically several magnitudes smaller than the susceptance S_{ij} , for the simplicity of mathematical treatment, we omit the $\cos(\cdot)$ term and only keep the $\sin(\cdot)$ term.

Let the rotor angles and the frequencies be denoted as $\boldsymbol{\theta} = [\theta_1 \ \cdots \ \theta_n]^\top$ and $\boldsymbol{\omega} = [\omega_1 \ \cdots \ \omega_n]^\top$, and the generator mechanical power injections be denoted as $\mathbf{P}_M = [P_{M_1} \ \cdots \ P_{M_n}]^\top$. The state-space representation of this nonlinear system is given by:

$$\begin{bmatrix} \dot{\boldsymbol{\theta}} \\ \dot{\boldsymbol{\omega}} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{Q}^{-1}\boldsymbol{\Lambda} & -\mathbf{Q}^{-1}\mathbf{K} \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \boldsymbol{\theta} \\ \boldsymbol{\omega} \end{bmatrix}}_{\mathbf{x}} + \underbrace{\begin{bmatrix} \mathbf{0} \\ \mathbf{Q}^{-1} \end{bmatrix}}_{\mathbf{B}} \mathbf{P}_M + \underbrace{\begin{bmatrix} \mathbf{0} \\ \mathbf{g}(\boldsymbol{\theta}) \end{bmatrix}}_{\mathbf{f}(\mathbf{x})}$$

where $\mathbf{g}(\boldsymbol{\theta}) = [g_1(\boldsymbol{\theta}) \ \cdots \ g_n(\boldsymbol{\theta})]^\top$ with

$$g_i(\boldsymbol{\theta}) = \sum_{j=1}^n A_{i+n,j} ((\theta_i - \theta_j) - \sin(\theta_i - \theta_j)),$$

where $A_{i+n,j}$ is the $(i+n, j)$ -th entry of \mathbf{A} , $\mathbf{Q} = \text{diag}(\{Q_i\}_{i=1}^n)$, $\mathbf{K} = \text{diag}(\{K_i\}_{i=1}^n)$, and $\boldsymbol{\Lambda}$ is a Laplacian matrix whose entries are specified in [4, Sec. IV-B]. For linearization (also known as DC approximation), the nonlinear part $\mathbf{g}(\mathbf{x})$ is assumed to be zero when the phase differences are small [4]. On the contrary, we deal with this term in the observer design directly. The information constraint is encoded in the observer matrix \mathbf{C} , which is a 6×20 matrix with its $(1, 1)$, $(2, 2)$, $(3, 3)$, $(4, 11)$, $(5, 12)$, $(6, 13)$ entries to be 1 and the remaining entries to be 0.

The nonlinearities of the system are due to $\mathbf{g}(\boldsymbol{\theta})$, which is a MIMO function of power phases. The phase differences are assumed to be bounded within $[-\Delta\theta_{\max}, \Delta\theta_{\max}]$ (i.e., $|\theta_i - \theta_j| \leq \Delta\theta_{\max}$ for $i, j \in [n]$), where θ_{\max} is typically less than $\pi/6$ under normal operations. The upper and lower bounds on partial derivatives $\frac{\partial}{\partial \theta_i} g_j(\boldsymbol{\theta})$ are given by³

$$\begin{cases} \underline{b}_{ij} = -A_{i+n,j} (1 - \cos(\Delta\theta_{\max})), \bar{b}_{ij} = 0 & \text{for } i \neq j \\ \underline{b}_{ij} = 0, \bar{b}_{ij} = \sum_{i \neq j} A_{i+n,j} (1 - \cos(\Delta\theta_{\max})) & \text{for } i = j \end{cases}$$

As listed in Table II, the proposed method (12) is able to find an asymptotically stable observer for $\Delta\theta_{\max}$ up to 0.92 (rad), covering a wide range of operational status. However, the LMI conditions in [13] can only certify observer stability for $\Delta\theta_{\max}$ up to 0.17.⁴ The method proposed in [9] is not applicable because it only deals with SISO nonlinear functions. Due to the scale of this system, the LPV approach proposed in [14] requires solving 2^{100} LMIs, which is computationally intractable.

³These bounds can be further refined by setting $A_{i+n,j} = 0$ in the calculation of bounds whenever $i \neq j$ and both i and j generators are directly observed. The effect of this refinement is left for future study.

⁴The Lipschitz constant could be estimate by the mean-value inequality

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \leq \sup_{0 \leq t \leq 1} \|\mathbf{J}_f(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))\|_2 \|\mathbf{x} - \mathbf{y}\|,$$

where $\mathbf{J}_f(\cdot)$ is the Jacobian of $\mathbf{f}(\cdot)$ and the supremum is taken over the line segment between \mathbf{x} and \mathbf{y} . This 2-norm can be further bounded using the matrix inequality $\|\mathbf{J}_f\|_2 \leq \sqrt{\|\mathbf{J}_f\|_1 \|\mathbf{J}_f\|_\infty}$, where the right-hand side can be bounded by the maximum phase difference. The Lipschitz constant γ_f can then be upper bounded by

$$\gamma_f \leq \sqrt{\|\mathbf{Q}^{-1}\boldsymbol{\Lambda}\|_1 \|\mathbf{Q}^{-1}\boldsymbol{\Lambda}\|_\infty (1 - \cos(\Delta\theta_{\max}))}.$$

TABLE II: Comparison results for different LMI methods for power system dynamic state estimation.

	LMI (50) in [13]	LMI (28) in [13]	LMI (35) in [14]	LMI (12)
$\Delta\theta_{\max}$ (rad)	< 0.17	< 0.17	†	> 0.92
computation time	< 0.5s	< 0.5s	†	< 3s

† At least 2^{100} 10×10 LMIs need to be solved, which is not implementable on a personal computer.

The tracking performance of our method is demonstrated in Fig. 3. To ensure the stability of the physical system, we deploy 10 independent neural-network controllers learned by reinforcement learning for \mathbf{P}_M , each of which can only observe states prescribed in Fig. 2 (details can be found in [28]). Because we assume that the inputs \mathbf{P}_M can be accessed directly, this does not change our design methodology in (12). The key advantage in the scalability of our multiplier-based approach can be verified in this real-world example.

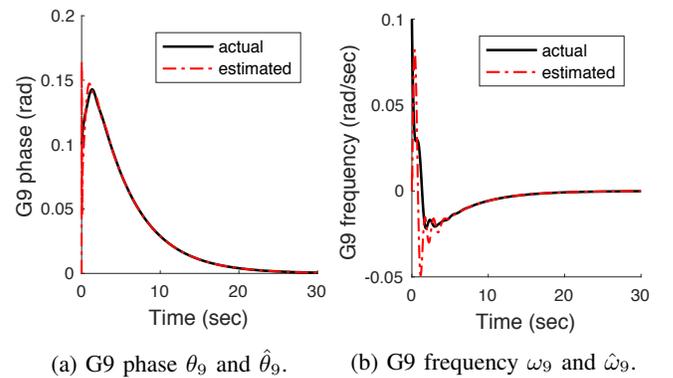


Fig. 3: Tracking performance of (12) for dynamic state estimation at the geographically separated generator G9. The observer is synthesized using (12) with $\Delta\theta_{\max} = \pi/4$ and $\rho = 8$ (\mathcal{L}_2 gain).

V. CONCLUSION

This study investigated an observer design methodology for nonlinear systems based on multipliers. Compared with existing methods, it can handle nonlinearities with far less conservatism, and can be employed for large-scale systems, as evaluated in a benchmark example and a case study on the dynamic power system state estimation. The multiplier-based framework can be generalized to incorporate multiple types of nonlinearities (e.g., time delays, time-varying and uncertain parameters), which significantly extends its applications to complex real-world problems.

REFERENCES

- [1] D. Simon, *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.
- [2] B. Açıkmeşe and M. Corless, “Observers for systems with nonlinearities satisfying incremental quadratic constraints,” *Automatica*, vol. 47, no. 7, pp. 1339–1348, 2011.
- [3] G. Phanomchoeng, R. Rajamani, and D. Piyabongkarn, “Nonlinear observer for bounded Jacobian systems, with applications to automotive slip angle estimation,” *IEEE Transactions on Automatic Control*, vol. 56, no. 5, pp. 1163–1170, 2011.

- [4] G. Fazelnia, R. Madani, A. Kalbat, and J. Lavaei, "Convex relaxation for optimal distributed control problems," *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 206–221, 2017.
- [5] A. J. Krener and W. Respondek, "Nonlinear observers with linearizable error dynamics," *SIAM Journal on Control and Optimization*, vol. 23, no. 2, pp. 197–216, 1985.
- [6] G. Ciccarella, M. Dalla Mora, and A. Germani, "A Luenberger-like observer for nonlinear systems," *International Journal of Control*, vol. 57, no. 3, pp. 537–556, 1993.
- [7] S. Raghavan and J. K. Hedrick, "Observer design for a class of nonlinear systems," *International Journal of Control*, vol. 59, no. 2, pp. 515–528, 1994.
- [8] R. Rajamani, "Observers for Lipschitz nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 43, no. 3, pp. 397–401, 1998.
- [9] M. Arcak and P. Kokotovic, "Observer-based control of systems with slope-restricted nonlinearities," *IEEE Transactions on Automatic Control*, vol. 46, no. 7, pp. 1146–1150, 2001.
- [10] X. Fan and M. Arcak, "Observer design for systems with multivariable monotone nonlinearities," *Systems & Control Letters*, vol. 50, no. 4, pp. 319–330, 2003.
- [11] F. Cacace, A. Germani, and C. Manes, "An observer for a class of nonlinear systems with time varying observation delay," *Systems & Control Letters*, vol. 59, no. 5, pp. 305–312, 2010.
- [12] M. Abbaszadeh and H. J. Marquez, "Nonlinear observer design for one-sided Lipschitz systems," in *Proc. of the American Control Conference*, 2010, pp. 5284–5289.
- [13] G. Phanomchoeng and R. Rajamani, "Observer design for Lipschitz nonlinear systems using Riccati equations," in *Proc. of the American Control Conference*, 2010, pp. 6060–6065.
- [14] A. Zemouche and M. Boutayeb, "On LMI conditions to design observers for Lipschitz nonlinear systems," *Automatica*, vol. 49, no. 2, pp. 585–591, 2013.
- [15] A. Alessandri and A. Rossi, "Increasing-gain observers for nonlinear systems: Stability and design," *Automatica*, vol. 57, pp. 180–188, 2015.
- [16] A. Megretski and A. Rantzer, "System analysis via integral quadratic constraints," *IEEE Transactions on Automatic Control*, vol. 42, no. 6, pp. 819–830, 1997.
- [17] J. C. Willems, "Dissipative dynamical systems part ii: Linear systems with quadratic supply rates," *Archive for rational mechanics and analysis*, vol. 45, no. 5, pp. 352–393, 1972.
- [18] G. Zames and P. Falb, "Stability conditions for systems with monotone and slope-restricted nonlinearities," *SIAM Journal on Control*, vol. 6, no. 1, pp. 89–108, 1968.
- [19] M. G. Safonov and V. V. Kulkarni, "Zames-Falb multipliers for MIMO nonlinearities," in *Proc. of the American Control Conference*, vol. 6, 2000, pp. 4144–4148.
- [20] M. Arcak and P. Kokotović, "Nonlinear observers: a circle criterion design and robustness analysis," *Automatica*, vol. 37, no. 12, pp. 1923–1930, 2001.
- [21] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*. SIAM, 1994, vol. 15.
- [22] A. Zemouche and M. Boutayeb, "A unified H_∞ adaptive observer synthesis method for a class of systems with both Lipschitz and monotone nonlinearities," *Systems & Control Letters*, vol. 58, no. 4, pp. 282–288, 2009.
- [23] A. Zemouche, R. Rajamani, H. Kheloufi, and F. Bedouhene, "Robust observer-based stabilization of Lipschitz nonlinear uncertain systems via LMIs-discussions and new design procedure," *International Journal of Robust and Nonlinear Control*, vol. 27, no. 11, pp. 1915–1939, 2017.
- [24] E. D. Sontag and Y. Wang, "Output-to-state stability and detectability of nonlinear systems," *Systems & Control Letters*, vol. 29, no. 5, pp. 279–290, 1997.
- [25] D. Angeli, "A Lyapunov approach to incremental stability properties," *IEEE Transactions on Automatic Control*, vol. 47, no. 3, pp. 410–421, 2002.
- [26] F. H. Clarke, *Optimization and nonsmooth analysis*. SIAM, 1990, vol. 5.
- [27] D. G. Luenberger, "Observing the state of a linear system," *IEEE Transactions on Military Electronics*, vol. 8, no. 2, pp. 74–80, 1964.
- [28] M. Jin and J. Lavaei, "Control-theoretic analysis of smoothness for stability-certified reinforcement learning," *Tech. report*. [Online]. Available: <http://www.ieor.berkeley.edu/~lavaei/RL.1.2018.pdf>

A. Proof of Lemma 1

Lemma 3 ([14]). *For a vector-valued function $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ that is differentiable with bounded partial derivatives, namely $\underline{b}_{ij} \leq \partial_j f_i(\mathbf{x}) \leq \bar{b}_{ij}$, there exist functions $\delta_{ij} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ with $\underline{b}_{ij} \leq \delta_{ij}(\mathbf{x}, \mathbf{z}) \leq \bar{b}_{ij}$ for all $i \in [m]$, $j \in [n]$, and $\mathbf{x}, \mathbf{z} \in \mathbb{R}^n$, such that $\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{z})$ can be written as*

$$\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{z}) = \sum_{i=1}^m \sum_{j=1}^n \delta_{ij}(\mathbf{x}, \mathbf{z}) \mathbf{e}_m^\top(i) \mathbf{e}_n^\top(j) (\mathbf{x} - \mathbf{z}), \quad (15)$$

where $\mathbf{e}_n(j)$ is an n -dimensional vector with its j -th component equal to 1 and the remaining ones equal to 0. \square

With the above result, we can prove the quadratic constraint in Lemma 1 for functions with bounded derivatives.

Proof of Lemma 1:

By Lemma 3, there exist functions $\delta_{ij} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ bounded by $\underline{b}_{ij} \leq \delta_{ij}(\mathbf{x}, \mathbf{y}) \leq \bar{b}_{ij}$, such that

$$\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y}) = \begin{bmatrix} \sum_{j=1}^n \delta_{1j}(\mathbf{x}, \mathbf{y})(x_j - y_j) \\ \vdots \\ \sum_{j=1}^n \delta_{mj}(\mathbf{x}, \mathbf{y})(x_j - y_j) \end{bmatrix}.$$

By defining $q_{ij} = \delta_{ij}(\mathbf{x}, \mathbf{y})(x_j - y_j)$, due to the bounds on partial derivatives, we have

$$(\delta_{ij}(\mathbf{x}, \mathbf{y}) - c_{ij})^2 \leq \bar{c}_{ij}^2,$$

for all $i \in [m]$ and $j \in [n]$, which can be reformulated as

$$\begin{bmatrix} x_j - y_j \\ q_{ij} \end{bmatrix}^\top \begin{bmatrix} \bar{c}_{ij}^2 - c_{ij}^2 & c_{ij} \\ c_{ij} & -1 \end{bmatrix} \begin{bmatrix} \star \\ \star \end{bmatrix} \geq 0.$$

The statement in Lemma 1 follows by introducing non-negative multipliers $\lambda_{ij} \geq 0$, and using the fact that $f_i(\mathbf{x}) - f_i(\mathbf{y}) = \sum_{j=1}^n q_{ij}$. \square

B. Proof of Theorem 2

By multiplying $\begin{bmatrix} \mathbf{e}^\top & \mathbf{q}^\top & \mathbf{v}^\top \end{bmatrix}^\top$ to the left and its transpose to the right of the augmented matrix in (12), and using the constraints $\mathbf{w} = \mathbf{W}\mathbf{q}$, the condition $\text{SDP}(\mathbf{P}, \mathbf{L}, \boldsymbol{\lambda}, \rho, \gamma)$ can be written as a dissipation inequality:

$$\dot{V}(\mathbf{e}) + \begin{bmatrix} \mathbf{e} \\ \mathbf{q} \end{bmatrix}^\top \mathbf{M}(\boldsymbol{\lambda}, \gamma) \begin{bmatrix} \mathbf{e} \\ \mathbf{q} \end{bmatrix} \leq \rho \mathbf{v}^\top \mathbf{v} - \frac{1}{\rho} \mathbf{e}^\top \mathbf{e},$$

where $V(\mathbf{e}) = \mathbf{e}^\top \mathbf{P} \mathbf{e}$ is known as the storage function, and $\dot{V}(\cdot)$ is its derivative with respect to time t . If $\text{SDP}(\mathbf{P}, \mathbf{L}, \boldsymbol{\lambda}, \rho, \gamma)$ is feasible with a solution $(\mathbf{P}, \mathbf{L}, \boldsymbol{\lambda}, \rho, \gamma)$, we have:

$$\dot{V}(\mathbf{e}) + \frac{1}{\rho} \mathbf{e}^\top \mathbf{e} - \rho \mathbf{v}^\top \mathbf{v} \leq 0, \quad (16)$$

which is satisfied at all times t . It can be integrated from $t = 0$ to $t = T$, and then it follows from $\mathbf{P} \succ 0$ that:

$$\int_0^T \|\mathbf{e}(t)\|^2 dt \leq \rho^2 \int_0^T \|\mathbf{v}(t)\|^2 dt + c, \quad (17)$$

where $c = \mathbf{e}(0)^\top \mathbf{P} \mathbf{e}(0)$ is a finite constant that depends on the initial tracking error. Hence, the observer is shown to be stable. \square