Automatica 79 (2017) 152–161

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

# Distributed moving horizon state estimation of two-time-scale nonlinear systems☆

Xunyuan Yin, Jinfeng Liu [1]

*Department of Chemical & Materials Engineering, University of Alberta, Edmonton, AB, T6G 1H9, Canada*

ARTICLE INFO

*Article history:*
Received 19 August 2015
Received in revised form
28 October 2016
Accepted 19 December 2016
Available online 6 March 2017

*Keywords:*
Large-scale systems
Moving horizon estimation
Observer-enhanced design
Time-scale multiplicity

ABSTRACT

In this paper, we focus on distributed moving horizon estimation (DMHE) for a class of two-time-scale nonlinear systems described in the framework of singularly perturbed systems. By taking advantage of the time-scale separation property, a two-time-scale system is first decomposed into a reduced-order fast system and a reduced-order slow system. The slow system is further decomposed into several interconnected slow subsystems. In the proposed distributed state estimation scheme, a local estimator is designed for each slow subsystem and for the reduced-order fast system. The slow subsystem estimators communicate with each other to exchange information and they are only required to send information to the fast system one-directionally. The fast system estimator does not send out any information. The local estimators are designed as observer-enhanced moving horizon estimators. Sufficient conditions on the convergence of the estimation error of the DMHE are derived. The application of the proposed DMHE to a chemical process example demonstrates its applicability and effectiveness.



## 1. Introduction

Complex and integrated systems are common occurrences in manufacturing industries (e.g., chemicals, petrochemicals and mineral processes). Model predictive control (MPC) systems are widely used in the manufacturing industries to ensure the quality of products while maximizing economic profits and guaranteeing operation safety as well as environmental sustainability. Due to the medium to large scales of many systems, the centralized control framework is not practical in terms of computational burden, organizational complexity, and fault tolerance (Christofides, Scattolini, Muñoz de la Peña, & Liu, 2013). The above considerations motivate significant research interests in distributed MPC (Christofides et al., 2013). While there are extensive results on distributed MPC, less attention has been given to distributed or decentralized state estimation which is equally important and is closely related to distributed control. It should be pointed out that there are some algorithms on decentralized or distributed Kalman filtering (e.g., Stanković, Stanković, & Stipanović, 2009). However, these algorithms typically do not account for system nonlinearity.

Recently, there are some results on distributed moving horizon state estimation (DMHE). Moving horizon estimation (MHE) is an online optimization-based technique and can handle nonlinearities, constraints and optimality considerations Alessandri, Baglietto, Battistelli, and Zavala (2010) and Haber and Verhaegen (2013). In Farina, Ferrari-Trecate, and Scattolini (2011), a DMHE algorithm was developed for nonlinear systems based on subsystem models. An observer-enhanced DMHE algorithm was developed in Zhang and Liu (2013a), where an auxiliary nonlinear observer is taken advantage of in the design of each local estimator. The auxiliary observer is used to calculate a reference state estimate based on which a confidence region is constructed every sampling time. Each local estimator optimizes its estimate within the confidence region. The observer-enhanced design is less sensitive to external noise compared with the auxiliary nonlinear observer. The convergence rate of the DMHE may be tuned by tuning the auxiliary observers. It was shown to be less dependent on the arrival cost, the estimation window size and have the potential to be used in output feedback control with provable closed-loop stability Zhang and Liu (2013a,b).

On the other hand, time-scale multiplicity is a common feature of many systems. For chemical processes, it usually arises due to the strong coupling of physicochemical phenomena (Han & Chung, 2001; Weekman & Nace, 1970). A direct application of

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Antonio Vicino under the direction of Editor Torsten Söderström.

*E-mail addresses:* xunyuan@ualberta.ca (X. Yin), jinfeng@ualberta.ca (J. Liu).

[1] Fax: +1 780 492 2881.

http://dx.doi.org/10.1016/j.automatica.2017.01.023

standard control or estimation methods without taking into account time-scale multiplicity to systems with different time scales may lead to ill-conditioning or even the loss of closed-loop stability (Christofides, 2000; Kokotovic, Khalil, & O'Reilly, 1986). The singular perturbation theory is the standard tool for the analysis of systems with time-scale multiplicity (Christofides, 2000; Kokotovic et al., 1986). Within the singular perturbation framework, the original system is typically decomposed into reduced-order subsystems with "fast" and "slow" dynamics. The majority of related results are on control system design for two-time-scale systems (e.g., Chen, Heidarinejad, Liu, Muñoz de la Peña & Christofides, 2011; Christofides, 2000; Kokotovic et al., 1986). Little attention has been given to state estimation of systems with time-scale multiplicity. In Nagy Kiss, Marx, Mourot, Schutz, and Ragot (2011), state estimation of a wastewater treatment plant was addressed via linearization in a centralized framework by neglecting the fast dynamics.

In this work, the scope is on the handling of time-scale multiplicity in state estimation. Specifically, we consider state estimation of a class of two-time-scale nonlinear systems. A system is first decomposed into a reduced-order fast system and several reduced-order slow subsystems. A fast MHE is designed for the fast system and a slow MHE is designed for each slow subsystem. The fast and slow MHEs form a DMHE scheme. Each MHE is designed via the observer-enhanced MHE technique (Liu, 2013). It is discovered that the slow MHEs are entirely decoupled from the fast MHE which is a significant difference from control of two-time-scale systems. The decoupling ensures that only one-directional information transmission from the slow MHEs to the fast MHE is needed and the fast MHE does not send out any information. Sufficient conditions are derived under which the proposed DMHE is guaranteed to give ultimately bounded estimation error under bounded system disturbances and measurement noise. The effectiveness of the proposed method is demonstrated via the application to a chemical process.

## 2. Preliminaries

### 2.1. Notation

The operator $|\cdot|$ denotes the Euclidean norm of a vector and $|\cdot|_Q^2$ represents the square of the weighted Euclidean norm of a vector, defined as $|x|_Q^2 = x^T Qx$ where $Q$ is a positive definite matrix. A function $f(x)$ is said to be Lipschitz with respect to its argument $x$ if there exists a positive constant $L_f^x$ such that $|f(x') - f(x'')| \leq L_f^x |x' - x''|$ holds for all $x'$ and $x''$ in a given region of $x$ and $L_f^x$ is the associated Lipschitz constant. A continuous function $\alpha : [0, a) \to [0, \infty)$ is said to belong to class $\mathcal{K}$ if it is strictly increasing and satisfies $\alpha(0) = 0$. A function $\beta(r, s)$ is said to be a class $\mathcal{KL}$ function if for each fixed $s$, $\beta(r, s)$ belongs to class $\mathcal{K}$ with respect to $r$, and for each fixed $r$, it is decreasing with respect to $s$, and $\beta(r, s) \to 0$ as $s \to \infty$. A function $f$ on an interval is said to be concave if for any $x$ and $y$ in the interval and for any $\alpha \in [0, 1]$, $f((1-\alpha)x + \alpha y) \geq (1-\alpha)f(x) + \alpha f(y)$. The symbol $\mathrm{diag}(v)$ denotes a diagonal matrix, in which the diagonal elements are the elements of vector $v$. The symbol $A^+$ denotes the pseudoinverse of a matrix (or vector) $A$. $\mathbb{I}$ denotes a set of integers defined as $\mathbb{I} = \{1, \ldots, m\}$.

### 2.2. System description

In this study, we consider a class of two-time-scale nonlinear systems that can be described in the framework of singularly perturbed systems as follows:

$$\dot{x}_s(t) = f(x_s(t), w_s(t), \epsilon) + \tilde{f}(x_s(t), x_f(t), \epsilon) \tag{1a}$$

$$\epsilon \dot{x}_f(t) = g(x_f(t), w_f(t), \epsilon) + \tilde{g}(x_s(t), x_f(t), \epsilon) \tag{1b}$$

$$y_s(t) = h_s(x_s(t)) + v_s(t) \tag{1c}$$

$$y_f(t) = h_f(x_f(t)) + v_f(t) \tag{1d}$$

where $x_s \in \mathbb{R}^{n_{x_s}}$ and $x_f \in \mathbb{R}^{n_{x_f}}$ are state vectors, $w_s \in \mathbb{R}^{n_{w_s}}$ and $w_f \in \mathbb{R}^{n_{w_f}}$ denote system disturbances, $y_s \in \mathbb{R}^{n_{y_s}}$ and $y_f \in \mathbb{R}^{n_{y_f}}$ are system outputs, $v_s \in \mathbb{R}^{n_{v_s}}$ and $v_f \in \mathbb{R}^{n_{v_f}}$ denote measurement noise, and $\epsilon$ is a small positive parameter reflecting the time-scale separation in the dynamics of the nonlinear system. The functions $f$ and $g$ depict, respectively, the dependence of the dynamics of $x_s$ and $x_f$ on themselves and associated system disturbances. The function $\tilde{f}$ characterizes the interaction between the dynamics of $x_s$ and the state vector $x_f$. Similarly, $\tilde{g}$ depicts the interaction between the dynamics of $x_f$ and $x_s$. It is assumed that functions $f$, $g$, $\tilde{f}$ and $\tilde{g}$ are all locally Lipschitz with respect to their arguments. Note that locally Lipschitz is a mild assumption on the continuity of the functions and it imposes limits on how fast the functions can change. The small parameter $\epsilon$ appears as a multiplier of the time derivative of state $x_f$, and the state $x_f$ evolves much faster than the state $x_s$ (Khalil, 2002). We will refer to $x_s$ as the slow states and $x_f$ as the fast states in the remainder. We assume that the measurements $y_s$ and $y_f$ are continuously available.

### 2.3. Two-time-scale decomposition

It is possible to decompose two-time-scale systems described in (1) into two separate reduced-order systems evolving in a fast and a slow time scales. This property will be taken advantage of in the design of the proposed distributed state estimation scheme.

First, we set $\epsilon = 0$ in (1) and obtain that:

$$\frac{dx_s(t)}{dt} = f(x_s(t), w_s(t), 0) + \tilde{f}(x_s(t), x_f(t), 0) \tag{2a}$$

$$0 = g(x_f(t), w_f(t), 0) + \tilde{g}(x_s(t), x_f(t), 0). \tag{2b}$$

We assume that there exists a unique isolated solution to the algebraic equation (2b):

$$x_f(t) = \hat{g}(x_s(t), w_f(t)) \tag{3}$$

for each pair of $(x_s, w_f)$, and the partial derivatives $\partial\hat{g}/\partial x$ and $\partial\hat{g}/\partial w$ are sufficiently smooth. This assumption is a standard one in two-time-scale decomposition and is used to ensure that $x_f$ can be uniquely expressed in terms of $x_s$ and $w_f$ (Kokotovic et al., 1986). Note that a control system is normally operated within an operating range and the assumption does not impose practical restrictions. Substituting (3) into (2a), the reduced-order slow system is obtained as follows:

$$\dot{\bar{x}}_s(t) = f(\bar{x}_s(t), w_s(t), 0) + \tilde{f}(\bar{x}_s(t), \hat{g}(\bar{x}_s(t), w_f(t)), 0). \tag{4}$$

Note that in (4), $\bar{x}_s$ is used to denote the state of the reduced-order slow system to indicate that the dynamics of the reduced-order slow system is (slightly) different from the dynamics of $x_s$ in the original system (1).

To derive the reduced-order fast system, we define a fast time scale $\tau = \frac{t}{\epsilon}$ and introduce the deviation variable $e_f := x_f - \hat{g}(x_s, w_f)$. The fast system (1b) can be rewritten in the following form: $\frac{de_f}{d\tau} = g\left(e_f + \hat{g}(x_s, w_f), w_f, \epsilon\right) + \tilde{g}\left(x_s, e_f + \hat{g}(x_s, w_f), \epsilon\right) - \epsilon\frac{\partial\hat{g}(x_s,w_f)}{\partial w_f}\dot{w}_f - \epsilon\frac{\partial\hat{g}}{\partial x_s}\left(f(x_s, w_s, \epsilon) + \tilde{f}(x_s, e_f + \hat{g}(x_s, w_f), \epsilon)\right)$. Setting $\epsilon$ to be zero and defining $G(e_f, x_s, w_f) := g\left(e_f + \hat{g}(x_s, w_f), w_f, 0\right) +$