Systems & Control Letters 115 (2018) 48–54

Contents lists available at ScienceDirect

Systems & Control Letters

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

Stochastic and adaptive optimal control of uncertain interconnected systems: A data-driven approach ☆

Tao Bian *, Zhong-Ping Jiang

Control and Networks Lab, Department of Electrical and Computer Engineering, Tandon School of Engineering, New York University, Brooklyn, NY 11201, USA

ARTICLE INFO

Article history:
Received 18 March 2016
Received in revised form 4 July 2017
Accepted 7 March 2018

Keywords:
Robust adaptive dynamic programming (RADP)
H^∞ control
Stochastic system
Zero-sum differential game
Value iteration
Small-gain

ABSTRACT

This paper provides a novel non-model-based, data-driven stochastic H^∞ control design for linear continuous-time stochastic interconnected systems with unknown dynamics. Our contributions are three-fold. First, we develop a tool to show how to assign an arbitrarily small input-to-output stochastic L^2 gain of the closed-loop system, by combining the gain assignment technique with the zero-sum dynamic game-based H^∞ control design. Second, robustness to dynamic uncertainties is tackled using the small-gain theory. Third, we develop a non-model-based stochastic robust adaptive dynamic programming (RADP) algorithm for adaptive optimal controller design. In sharp contrast to the existing methods, the obtained algorithm is based on value iteration (VI), and the knowledge of an initial stabilizing control policy is no longer needed. An example of a power electronic system is adopted to illustrate the obtained results.

1. Introduction

H^∞ control is a fundamentally important topic in control theory; see, [1], and references therein. Among the research devoted to H^∞ control, a large amount of attention has been paid to the stochastic H^∞ control problem [2–5].

One way of finding a stochastic H^∞ controller is by solving a zero-sum dynamic game [4,1]. For linear systems, the H^∞ controller is obtained through solving the generalized algebraic Riccati equation (GARE). However, there are several unsolved difficulties within the existing stochastic differential game framework. First, existing stochastic H^∞ controllers are devised based on the given stochastic L^2 gain (for example, the gain is specifically picked as one in [3, Section 3]), and there is no efficient way to assign this L^2 gain to be arbitrarily small, without requiring restrictive assumptions. Second, solving the GARE directly requires the precise knowledge of the system model, and it is still an open problem how to conduct a data-driven design to find the stochastic H^∞ controller in the absence of precise model knowledge.

This paper has three major contributions. First, we develop a new tool to assign an arbitrarily small input-to-output stochastic L^2 gain for a class of partially linear interconnected systems, by combining the nonlinear gain assignment technique [6–9] with the zero-sum stochastic differential game approach [1,4]. Second, the dynamic uncertainty is considered in our model. Different from the past literature, we investigate the influence of dynamic uncertainties on the stochastic H^∞ gain of the closed-loop system. Third, we conduct a new data-driven H^∞ control design with unknown model knowledge. The main advantage of the data-driven approach is that the system matrices are not necessarily known, except assuming some bounds on these matrices. Moreover, different from existing learning-based methods [10–22], our method is based on a continuous-time stochastic VI, and a stabilizing initial control policy is no longer required to start the learning process. Rigorous convergence and stability analysis is presented for the proposed algorithm and the closed-loop system.

The remainder of this paper is organized as follows. The problem formulation is given in Section 2. Section 3 provides a partial-state feedback H^∞ control design with concrete robust stability analysis. In Section 4, a data-driven online learning algorithm is presented, along with convergence and stability analysis. An example of a power electronic system is given in Section 5. Finally, the conclusion is drawn in Section 6.

Notation: In this paper, $\mathscr{S}^n$ denotes the space of all n-by-n real symmetric matrices. $\mathscr{S}^n_+ = \{P \in \mathscr{S}^n : P \geq 0\}$. For any $A \in \mathbb{R}^{n\times m}$, $\mathrm{vec}(A) = [a_1^T, a_2^T, \ldots, a_m^T]^T$, where $a_i \in \mathbb{R}^n$ is the ith column of A. For any $A \in \mathscr{S}^n$, $\mathrm{vecs}(A) = [a_{11}, a_{12}, \ldots, a_{1n}, a_{22}, a_{23}, \ldots, a_{n-1n}, a_{nn}]^T$, where $a_{ij} \in \mathbb{R}$ is the (i, j)th element of A. Denote $w = [w_1, w_2, \ldots, w_q]^T$ as a q-dimensional standard Brownian motion, and $\mathscr{F}_t$ as the σ-field generated by $w(s)$, $0 \leq s \leq t$.

☆ This work has been partially supported by the National Science Foundation under Grants ECCS-1230040 and ECCS-1501044.

* Corresponding author.
E-mail address: tbian@nyu.edu (T. Bian).

https://doi.org/10.1016/j.sysconle.2018.03.001
0167-6911/© 2018 Elsevier B.V. All rights reserved.

2. Problem formulation

Consider the following class of stochastic interconnected systems:

$$dx = Axdt + B\left((z + \Delta_1(\varsigma, x, v))dt + \sum_{i=1}^{q} A_{0i}xdw_i\right), \tag{1}$$

$$dz = G\left((u + \Delta_2(\varsigma, x, z, v))dt + \sum_{i=1}^{q}(E_{0i}x + F_{0i}z)dw_i\right), \tag{2}$$

$$d\varsigma = f_0(\varsigma, x)dt + \sum_{i=1}^{q} g_i(\varsigma, x)dw_i, \tag{3}$$

$$y = x, \tag{4}$$

where $(x, z, \varsigma) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ is the state; $u : \mathbb{R}_+ \to \mathbb{R}^m$, and $v : \mathbb{R}_+ \to \mathbb{R}^{m_v}$ are $\{\mathscr{F}_t\}$-adapted random processes representing the control input and the external disturbance of the system, respectively; $y \in \mathbb{R}^q$ is the output; $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $G \in \mathbb{R}^{m\times m}$, $A_{0i} \in \mathbb{R}^{m\times n}$, $E_{0i} \in \mathbb{R}^{m\times n}$, and $F_{0i} \in \mathbb{R}^{m\times m}$, $i = 1, 2, \ldots, q$, are system matrices with G nonsingular; $f_0 : \mathbb{R}^p\times\mathbb{R}^n \to \mathbb{R}^p$, $g_i : \mathbb{R}^p\times\mathbb{R}^n \to \mathbb{R}^p$, $\Delta_1 : \mathbb{R}^p \times \mathbb{R}^n \times \mathbb{R}^{m_v} \to \mathbb{R}^m$, and $\Delta_2 : \mathbb{R}^p \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m_v} \to \mathbb{R}^m$ are unknown locally Lipschitz functions satisfying $f_0(0, 0) = 0$, $g_i(0, 0) = 0$, $\Delta_1(0, 0, 0) = 0$, $\Delta_2(0, 0, 0, 0) = 0$. Suppose system (1) is mean square stabilizable, with z as the input and $\Delta_1 = 0$. The ς-subsystem (3) is often referred to as the dynamic uncertainty [8]. Related forms of system (1)–(4) have been studied in [23–26].

Suppose the dynamic uncertainty (3) is unknown, and we only know some bounds on the system matrices in (1)–(2). Before presenting our problem formulation, we first give the following definition:

Definition 1. We say a control design method is *data-driven*, if the controller is designed directly using input–output or input-state data of the controlled system, without using or identifying the explicit knowledge of the mathematical model of the controlled process.

In this paper, we aim at solving the following problem:

Problem 1. Given an arbitrary $\gamma > 0$, find a partial-state feedback law $u = \mu(x, z)$, that does not depend on the state ς of the dynamic uncertainty (3), via data-driven adaptive optimal control design, such that the closed-loop system (1)–(4) (a) is exponentially stable at the origin in the mean square sense when $v \equiv 0$; and (b) admits a linear stochastic L^2 gain from v to y less than or equal to γ.

Two obstacles prohibit us from solving Problem 1 directly via the standard stochastic H^∞ control technique [2–4]. First, due to the presence of the dynamic uncertainty and the unmatched disturbance input, the zero-sum differential game approach is not sufficient to freely assign the stochastic L^2 gain. Second, it is still an open problem how to develop a data-driven robust adaptive optimal control design to solve the stochastic H^∞ control problem.

In this paper, we solve Problem 1 in two steps. First, assuming all the matrices in system (1)–(2) are fully accessible, we use backstepping to study the adaptive/optimal control for a cascade of subsystems. The gain of each subsystem is then assigned by solving a stage H^∞ optimization problem. This result is presented in Theorem 1. Then, we develop a data-driven approach to approximate the stochastic H^∞ controller via online learning. This is done in Section 4.

3. Stochastic robust optimal control design

First, we focus on the nominal system (1)–(2), and derive the stochastic L^2 gain from Δ to y. Consider (1) with z regarded as the virtual control input and Δ_1 as an external disturbance input. Write the cost corresponding to (1) as

$$\mathcal{J}_1(x(0); z, \Delta_1) = \mathcal{E}\left[\int_0^\infty (x^T Q_0 x + \lambda_1|z|^2 - \gamma_1^2|\Delta_1|^2)dt\right],$$

where $\gamma_1 > 0$, $\gamma_1^2 > \lambda_1 > 0$, and $Q_0 = Q_0^T > 0$. Then, from stochastic differential game theory, the pair of stochastic H^∞ controller and the worst case disturbance, i.e., (z^*, Δ_1^*), is solved as

$$z^* = -\lambda_1^{-1}B^T P^* x \equiv -K^*x, \quad \Delta_1^* = \gamma_1^{-2}B^T P^* x,$$

where $P^* = P^{*T} > 0$ is the solution to

$$0 = A^T P + PA - (\lambda_1^{-1} - \gamma_1^{-2})PBB^T P + \Pi_1(P) + Q_0, \tag{5}$$

and $\Pi_1(P) = \sum_{i=1}^{q} A_{0i}^T B^T PBA_{0i}$. Note that $\lambda_1^{-1} - \gamma_1^{-2} > 0$ by our definition. Since system (1) is mean square stabilizable with $H_1 = 0$ and z as the input, P^* is always well defined by [27, Corollary (11.4.14)].

Now define $\xi = z - z^*$. From (1) and (2), we have

$$d\xi = \bar{F}\xi dt + \bar{E}xdt + G\left((u + \bar{\Delta}_2)dt + \sum_{i=1}^{q}(\bar{E}_{0i}x + F_{0i}\xi)dw_i\right), \tag{6}$$

where $\bar{F} = K^*B$, $\bar{E} = K^*A - K^*BK^*$, $\bar{E}_{0i} = E_{0i} + G^{-1}K^*BA_{0i} - F_{0i}K^*$, and $\bar{\Delta}_2 = \Delta_2 + G^{-1}K^*B\Delta_1$.

Consider system (6) with $\bar{E}_{0i} = \bar{E} = 0$ and the cost

$$\mathcal{J}_2(\xi(0); u, \bar{\Delta}_2) = \mathcal{E}\left[\int_0^\infty (\xi^T W\xi + \lambda_2|u|^2 - \gamma_2^2|\bar{\Delta}_2|^2)dt\right],$$

where $\gamma_2 > 0$, $\gamma_2^2 > \lambda_2 > 0$, and $W = W^T > 0$. Then, similar to the case for the x-system, the optimal solution pair $(u^*, \bar{\Delta}_2^*)$ is obtained as

$$u^* = -\lambda_2^{-1}G^T M^*\xi \equiv -L^*\xi, \quad \bar{\Delta}_2^* = \gamma_2^{-2}G^T M^*\xi,$$

where $M^* = M^{*T} > 0$ is the solution to

$$0 = \bar{F}^T M + M\bar{F} - (\lambda_2^{-1} - \gamma_2^{-2})MGG^T M + \Pi_2(M) + W, \tag{7}$$

and $\Pi_2(M) = \sum_{i=1}^{q} F_{0i}^T G^T MGF_{0i}$. Note that since $\lambda_2^{-1} - \gamma_2^{-2} > 0$ and G is nonsingular, M^* is also well defined by [27, Corollary (11.4.14)].

Now, we rewrite system (1) and (6) in a compact form:

$$d\zeta = A_D\zeta dt + B_D udt + B_\Delta\left(\bar{\Delta}dt + \sum_{i=1}^{q} A_{wi}\zeta dw_i\right), \tag{8}$$

where $\zeta = [x^T, \xi^T]^T$, $\bar{\Delta} = [\Delta_1^T, \bar{\Delta}_2^T]^T$,

$$A_D = \begin{bmatrix} A - BK^* & B \\ \bar{E} & \bar{F} \end{bmatrix}, \quad B_D = \begin{bmatrix} 0 \\ G \end{bmatrix}, \quad B_\Delta = \begin{bmatrix} B & 0 \\ 0 & G \end{bmatrix},$$

$$A_{wi} = \begin{bmatrix} A_{0i} & 0 \\ \bar{E}_{0i} & F_{0i} \end{bmatrix}.$$

Lemma 1. *For any $\gamma_1, \gamma_2 > 0$, there exist Q_0, λ_1, λ_2, and W, such that system (8) under $u = u^*$ admits a finite stochastic L^2 gain less than or equal to $\max\{\gamma_1, \gamma_2\}$, with $\bar{\Delta}$ as the input and ζ as the output.*

Proof. First, given any $\gamma_1 > 0$ and $\gamma_2 > 0$, we can pick $Q_0 = Q_0^T \geq I_n$, $0 < \lambda_1 < \gamma_1^2$, and thus fix P^*. Note from [28] that for any $W = W^T > 0$, $\lim_{\lambda_2\to 0} M^*$ is well defined. Since G is nonsingular, one easily has by (7) that $\lim_{\lambda_2\to 0}|M^*| = 0$. Hence, we can pick $W > I_m$ and a sufficiently small $0 < \lambda_2 < \gamma_2^2$, such that

$$Q := \begin{bmatrix} Q_0 + \lambda_1 K^{*T}K^* - \Pi_3(M^*) & -P^*B - \bar{E}^T M^* - \Pi_4(M^*) \\ -B^T P^* - M^*\bar{E} - \Pi_4^T(M^*) & W \end{bmatrix}$$
$$\geq I_{n+m},$$