Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Learning Koopman Operator under Dissipativity Constraints

Keita Hara *, **Masaki Inoue** *, **Noboru Sebe** **

* *Department of Applied Physics and Physico-Informatics, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa, Japan.*
** *Department of Artificial Intelligence, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka, Japan.*

Abstract: This paper addresses a learning problem for nonlinear dynamical systems with incorporating any specified dissipativity property. The nonlinear systems are described by the Koopman operator, which is a linear operator defined on the infinite-dimensional lifted state space. The problem of learning the Koopman operator under specified quadratic dissipativity constraints is formulated and addressed. The learning problem is in a class of the non-convex optimization problem due to nonlinear constraints and is numerically intractable. By applying the change of variable technique and the convex overbounding approximation, the problem is reduced to sequential convex optimization and is solved in a numerically efficient manner. Finally, a numerical simulation is given, where high modeling accuracy achieved by the proposed approach including the specified dissipativity is demonstrated.

Keywords: Learning, Dissipativity, Koopman Operator, Linear Matrix Inequality

1. INTRODUCTION

Artificial neural networks have re-developed recently and have brought high modeling accuracy in data-driven modeling of complex nonlinear systems. In the development, multi-layered hierarchical models play a central role, and their efficient learning methods receive considerable attention in various areas (see e.g., the works by De la Rosa and Yu (2016), Jin et al. (2016)). For example, De la Rosa and Yu (2016) presents an identification method for nonlinear dynamical systems by using deep learning techniques. Jin et al. (2016) proposes a deep reconstruction model (DRM) to analyze the characteristics of nonlinear systems.

The drawback of such learning methods, in particular for researchers or engineers, is the gap between the constructed model and some *a priori* information on a physical system. Although the model may fit a given data-set and emulate the system behavior accurately, it cannot necessarily possess some practically essential properties of the system. If we know *a priori* information on the properties, it is natural to try to incorporate them into the model. For *linear* dynamical systems, a variety of learning methods that incorporate *a priori* information have been studied well. For example, the subspace identification method is combined with the *a priori* information including stability by Lacy and Bernstein (2003), eigenvalue location by Okada and Sugie (1996); Miller and De Callafon (2013), steady-state property by Alenany et al. (2011); Yoshimura et al. (2019), moments by Inoue (2019), and positive-realness studied by Goethals et al. (2003); Hoagg et al. (2004), more general frequency-domain property by Abe et al. (2016). However, to the best of the authors' knowledge, for *nonlinear* dynamical systems, learning methods with *a priori* information have not been studied well.

This paper addresses a learning problem for nonlinear dynamical systems with incorporating the *a priori* information on *dissipativity*, which is proposed by Willems (1972) and developed e.g., by Hill and Moylan (1976, 1977). To this end, the *nonlinear* system is described with the *Koopman operator*, which is a *linear* operator defined on infinite-dimensional lifted state space and has been applied to analysis of nonlinear dynamical systems. See e.g., the works by Koopman (1931), Williams et al. (2015), Korda and Mezić (2018). Then, the learning problem is reduced to the data-driven finite-dimensional approximation of the Koopman operator onto the dissipativity constraint.

The approximation is formulated as the minimization of a convex cost function, which measures the consistency of the model and data, subject to a nonlinear matrix inequality, which represents the dissipation inequality. The formulated problem, which is called Problem 1, is in a class of the non-convex optimization and is numerically intractable. Therefore, we aim at approximating Problem 1 to derive a numerically efficient algorithm that is composed of the solutions to the following two problems. 1) By applying the change of variable technique to the nonlinear matrix inequality, we derive a linear matrix inequality (LMI) constraint. In addition, the approximation of the cost function reduces Problem 1 to a convex optimization, which is called Problem 2 and provides the feasible solution to Problem 1. 2) The convex overbounding approximation method proposed by Sebe (2018) is applied to the nonlinear matrix inequality of Problem 1 to derive its inner approximation. Then, the derived convex optimization problem, which is called Problem 3, is sequentially solved by starting from the initial guess obtained in Problem 2. It is guaranteed that the overall algorithm generates a less conservative solution than the solution obtained in Problem 2.

The remaining parts of this paper are organized as follows. In Section 2, we review theory of the Koopman operator and dissipativity. In Section 3, the problem of learning the Koopman operator with the dissipativity-constraint is formulated, and the learning algorithm is proposed. In Section 4, a numerical simulation is performed, in which the learning problem of a dissipative nonlinear system is addressed and the effectiveness of the algorithm is presented. Section 5 concludes the works in this paper.

2. PRELIMINARIES

2.1 Koopman Operator Theory

In this section, we review *Koopman operator theory* to show its application to control systems based on the work by Korda and Mezić (2018).

Koopman Operator We consider a discrete-time nonlinear system described by

$$\begin{cases} x(k+1) &= f\left(x(k), u(k)\right), \\ y(k) &= g\left(x(k)\right), \end{cases} \tag{1}$$

where k is the discrete time, $u \in \mathbb{R}^m$ is the input, $x \in \mathbb{R}^n$ is the state, $y \in \mathbb{R}^l$ is the output, and $f(x,u) : \mathbb{R}^{n+m} \to \mathbb{R}^n$ and $g(x) : \mathbb{R}^n \to \mathbb{R}^l$ are the nonlinear functions. Let z denote the extended state

$$z := \begin{bmatrix} x \\ u \end{bmatrix} \in \mathbb{R}^{n+m}.$$

Further let $\mathcal{F}$ be the nonlinear operator defined by

$$\mathcal{F}(z) := \begin{bmatrix} f(x,u) \\ \mathcal{S}(u) \end{bmatrix},$$

where $\mathcal{S}$ is the time-shift operator defined as

$$\mathcal{S}(u(k)) := u(k+1).$$

Then, the time evolution of z is described by

$$z(k+1) = \mathcal{F}(z(k)). \tag{2}$$

Now, we let $\phi_{\inf}(z)$ denote the infinite-dimensional lifting function described by

$$\phi_{\inf}(z) = \begin{bmatrix} \phi_1(z) \\ \phi_2(z) \\ \vdots \end{bmatrix}. \tag{3}$$

Here, we introduce the Koopman operator $\mathcal{K}$ as

$$\mathcal{K}(\phi_{\inf}(z)) := \phi_{\inf}(\mathcal{F}(z)).$$

Then, the time evolution of $\phi_{\inf}(z)$ is described by

$$\phi_{\inf}(z(k+1)) = \mathcal{K}(\phi_{\inf}(z(k))). \tag{4}$$

Note that the Koopman operator is a linear operator defined on the infinite-dimensional state space, while expressing nonlinear dynamical systems (see e.g., the work by Korda and Mezić (2018)).

Approximation of Koopman Operator Since the Koopman operator is the infinite-dimensional operator, it is difficult to be handled in numerical calculations. In this subsection, we give the finite-dimensional approximation of the Koopman operator. To this end, we define the N_ϕ-dimensional lifting function $\phi(z) : \mathbb{R}^{n+m} \to \mathbb{R}^{N_\phi}$ as

$$\phi(z) = \begin{bmatrix} \phi_1(z) \\ \vdots \\ \phi_{N_\phi}(z) \end{bmatrix} \in \mathbb{R}^{N_\phi}.$$

Furthermore, we let $\mathcal{A} \in \mathbb{R}^{N_\phi \times N_\phi}$ be a finite-dimensional matrix that approximates the Koopman operator $\mathcal{K}$, i.e., the error

$$\|\mathcal{A}\phi(z) - \phi(\mathcal{F}(z))\| \tag{5}$$

is sufficiently small. With this $\mathcal{A}$, we have the expression

$$\phi(z(k+1)) \approx \mathcal{A}\phi(z(k)), \tag{6}$$

which approximately describes the behavior of $\phi_{\inf}(z)$, defined by (4).

In this paper we propose the method of the data-driven approximation of $\mathcal{K}$, i.e., the learning method of $\mathcal{A}$ by using some data. In the method, we aim at constructing the model (6) that is *compatible* with controller design. It is tractable for controller design and its implementation that the model is linear to the input u. To this end, we further specialize the class of the lifting function $\phi(z)$ in the following form

$$\phi(z) = \begin{bmatrix} \psi(x) \\ u \end{bmatrix} \in \mathbb{R}^{N+m},$$

where $\psi(x) : \mathbb{R}^n \to \mathbb{R}^N$ is N-dimensional lifting function given by

$$\psi(x) = \begin{bmatrix} \psi_1(x) \\ \vdots \\ \psi_N(x) \end{bmatrix} \in \mathbb{R}^N,$$

and $N + m = N_\phi$ holds. Let the matrix $\mathcal{A}$ be partitioned as

$$\mathcal{A} = \begin{bmatrix} A & B \\ * & * \end{bmatrix} \in \mathbb{R}^{(N+m)\times(N+m)},$$

where $A \in \mathbb{R}^{N\times N}$ and $B \in \mathbb{R}^{N\times m}$. Then, it follows from (6) that the following expression of the time-evolution of $\psi(x)$ holds.

$$\psi(x(k+1)) \approx A\psi(x(k)) + Bu(k).$$

In addition, we give the approximation of the output equation in (1) by

$$y(k) \approx C\psi(x(k)),$$

where $C \in \mathbb{R}^{l\times N}$. For simplicity of notation, we let $\psi(k) = \psi(x(k))$. Then, we obtain the state-space model defined on the functional space as

$$\begin{cases} \psi(k+1) &= A\psi(k) + Bu(k), \\ y(k) &= C\psi(k). \end{cases} \tag{7}$$

The model (7) approximately describes the nonlinear input-output behavior generated by (1). In this paper, the model (7) is called the "Koopman model". The aim of this paper is to propose the learning method of the system matrices (A, B, C) based on some data-sets.

Learning Koopman Operator For simplicity of notation, we define the following data-matrices based on the sequences of the input, output, and state of the system (1).

$$\begin{aligned} U_k &:= [u(k)\ u(k+1)\ \cdots\ u(k+M-1)] \in \mathbb{R}^{m\times M}, \\ Y_k &:= [y(k)\ y(k+1)\ \cdots\ y(k+M-1)] \in \mathbb{R}^{l\times M}, \\ \Psi_k &:= [\psi(k)\ \psi(k+1)\ \cdots\ \psi(k+M-1)] \in \mathbb{R}^{N\times M}, \\ \Psi_{k+1} &:= [\psi(k+1)\ \psi(k+2)\ \cdots\ \psi(k+M)] \in \mathbb{R}^{N\times M}. \end{aligned}$$

It should be noted that Ψ_k and Ψ_{k+1} are constructed by using the measured data on the state, $\{x(k), \ldots, x(k+M)\}$. In this paper, it is assumed that the data-set $(U_k, Y_k, \Psi_k, \Psi_{k+1})$ is given and available for learning the Koopman operator that expresses (1). The problem of learning is formulated as follows.

Given the data-matrices $(U_k, Y_k, \Psi_k, \Psi_{k+1})$, solve the optimization problem:

$$\min_{A,B,C} \quad J_1(A,B) + J_2(C), \tag{8}$$

where $J_1(A,B)$ and $J_2(C)$ are given by

$$J_1(A,B) := \left\| \Psi_{k+1} - [A\ B] \begin{bmatrix} \Psi_k \\ U_k \end{bmatrix} \right\|_F^2, \tag{9}$$

$$J_2(C) := \|Y_k - C\Psi_k\|_F^2 . \tag{10}$$

The solution to the optimization problem (8) provides the system matrices $(A^\dagger, B^\dagger, C^\dagger)$ of the Koopman model (7). It is assumed that $\left[\Psi_k^\top U_k^\top\right]^\top$ is of full row rank, which is a natural assumption when rich data is available for learning. Then, the learned matrices $(A^\dagger, B^\dagger, C^\dagger)$ are uniquely determined by any given data.

2.2 Dissipativity

In this subsection, we review *dissipativity* of dynamical systems. Dissipativity is a property of characterizing dynamical systems and plays an important role in system analysis, in particular, the analysis of feedback or more general interconnection of dynamical systems (see e.g., the pioneering work by Willems (1972) and developments e.g., by Hill and Moylan (1976, 1977)). Dissipativity is defined for the input-output system (1) as follows.

Definition 1. Given a scalar function $s(u,y) : \mathbb{R}^{m+l} \to \mathbb{R}$, the system (1) is said to be *dissipative* for $s(u,y)$ if there is a non-negative function $V(x) : \mathbb{R}^n \to \mathbb{R}_+$ such that the inequality

$$V(x(k)) - V(x(0)) \leq \sum_{\tau=0}^{k} s(u(\tau), y(\tau)), \quad \forall k \geq 0 \tag{11}$$

holds.

The functions $s(u,y)$ and $V(x)$ and the inequality (11) are called the supply rate, storage function, and dissipation inequality, respectively.

A characterization of dissipative *linear* dynamical systems is given. We specialize the supply rate $s(u,y)$ in the quadratic form as

$$s(u,y) = -\begin{bmatrix} y \\ u \end{bmatrix}^{\mathrm{T}} \Xi \begin{bmatrix} y \\ u \end{bmatrix}, \tag{12}$$

where Ξ is the real symmetric matrix of

$$\Xi = \begin{bmatrix} \Xi_{11} & \Xi_{12} \\ \Xi_{12}^{\mathrm{T}} & \Xi_{22} \end{bmatrix}. \tag{13}$$

Even for the specialization of the supply rate, the dissipativity includes some important property of dynamical systems. For example, the dissipativity with respect to $(\Xi_{11}, \Xi_{12}, \Xi_{22}) = (0, -1, 0)$ represents the passivity of dynamical systems, and that with respect to $(\Xi_{11}, \Xi_{12}, \Xi_{22}) = (1, 0, -\gamma)$ for some positive constant γ represents the bounded L_2 gain.

The dissipativity of linear input-output systems, e.g., described by (7), is characterized by the following lemma.

Lemma 1. The following statements (i) and (ii) are equivalent (see the book by Brogliato et al. (2007)).

(i) The Koopman model (7) is dissipative for the supply rate $s(u,y)$ of (12).

(ii) There exists a symmetric matrix P such that the following inequalities hold.

$$P \succ 0 \tag{14}$$

$$\begin{bmatrix} A & B \\ I & 0 \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} P & 0 \\ 0 & -P \end{bmatrix} \begin{bmatrix} A & B \\ I & 0 \end{bmatrix} + \begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix}^{\mathrm{T}} \Xi \begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix} \prec 0. \tag{15}$$

3. LEARNING KOOPMAN OPERATOR WITH DISSIPATIVITY-CONSTRAINTS

In this section, we propose a learning method of nonlinear dynamical systems with incorporating *a priori* information on the system dissipativity. We assume that the supply rate $s(u,y)$ characterizing the system dissipativity is already given and available for learning. Then, we aim at incorporating the dissipativity information into the Koopman model (7).

3.1 Problem Setting

We aim at constructing the Koopman model (7) that satisfies the dissipation inequality (11) based on some data-sets. This learning problem of the system matrices (A, B, C) of (7) is reduced to the problem of (8) subject to the dissipativity constraints (14) and (15). The problem is mathematically formulated as follows.

Problem 1. Given the real symmetric matrix Ξ and the data-matrices $(U_k, Y_k, \Psi_k, \Psi_{k+1})$, solve the optimization problem:

$$\min_{P,A,B,C} \quad J_1(A,B) + J_2(C)$$
$$\text{sub to} \quad (14), (15).$$

Suppose that the optimal solution to Problem 1 is given by (P^*, A^*, B^*, C^*). Then the Koopman model (7) with (A^*, B^*, C^*) is dissipative for the supply rate $s(u,y)$ of (12).

Note that the dissipativity constraint, described by (14) and (15), is non-convex in decision variables (P, A, B, C). It is not numerically tractable to solve Problem 1. In the next subsection, we try to approximately reduce the problem to a convex one.

3.2 Convex Approximation of Problem 1

On the basis of the variable transformation technique by Hoagg et al. (2004); Abe et al. (2016), the nonlinear inequality (15) is reduced into a linear matrix one. We expand (15) as

$$\begin{bmatrix} P - C^{\mathrm{T}}\Xi_{11}C & -C^{\mathrm{T}}\Xi_{12} \\ -\Xi_{12}^{\mathrm{T}}C & -\Xi_{22} \end{bmatrix} - [A\ B]^{\mathrm{T}} P\, [A\ B] \succ 0. \tag{16}$$

Then, we apply the variable transformation to (P, A, B). We let

$$R = PA, \quad S = PB \tag{17}$$

to reduce (16) to the inequality

$$\begin{bmatrix} P - C^{\mathrm{T}}\Xi_{11}C & -C^{\mathrm{T}}\Xi_{12} & R^{\mathrm{T}} \\ -\Xi_{12}^{\mathrm{T}}C & -\Xi_{22} & S^{\mathrm{T}} \\ R & S & P \end{bmatrix} \succ 0. \tag{18}$$

Now, suppose that C is given, e.g., by just minimizing $J_2(C)$ based on the data-set (Y_k, Φ_k). Then, the inequality (18) is linear in the matrices (P, R, S), which is numerically tractable.

There is the drawback in the variable transformation of (17): the cost function $J_1(A,B)$ of (9) becomes non-convex in the transformed variables (P, R, S), which is numerically intractable. To overcome the drawback and to numerically obtain the feasible solution to Problem 1, we approximately transform $J_1(A,B)$ to a convex one. To this end, we introduce $W = P$, as a weighting matrix, into $J_1(A,B)$ to define a new cost function as follows.

$$J_{1,W}(P,R,S) = \left\| W \left(\Psi_{k+1} - [A\ B] \begin{bmatrix} \Psi_k \\ U_k \end{bmatrix} \right) \right\|_F^2 = \left\| P\Psi_{k+1} - [R\ S] \begin{bmatrix} \Psi_k \\ U_k \end{bmatrix} \right\|_F^2. \quad (19)$$

The function $J_{1,W}(P,R,S)$ of (19) is convex in the matrices (P,R,S). The minimization problem of $J_{1,W}(P,R,S)$ under the inequalities (14) and (18) is in the class of the convex optimization. The optimization problem is summarized as follows.

Problem 2. Given the system matrix C, the real symmetric matrix Ξ, and the data-matrices $(U_k, \Psi_k, \Psi_{k+1})$, solve the optimization problem:

$$\min_{P,R,S} \quad J_{1,W}(P,R,S)$$
$$\text{sub to} \quad (14), (18).$$

Suppose that Problem 2 is feasible and that the solution $(\hat{P}, \hat{R}, \hat{S})$ is given. Then, we obtain the system matrices as

$$\hat{A} = \hat{P}^{-1}\hat{R}, \quad \hat{B} = \hat{P}^{-1}\hat{S}. \quad (20)$$

We have the following proposition for the constructed model based on the solution to Problem 2.

Proposition 1. Suppose that Problem 2 is feasible and that system matrices are given by (20). Then, the quadruplet $(\hat{P}, \hat{A}, \hat{B}, C)$ is the *feasible* solution to Problem 1. In other words, the Koopman model (7) with $(\hat{A}, \hat{B}, C)$ is dissipative for the supply rate $s(u,y)$ of (12).

As implied in the proposition, the solution to Problem 2 is the *feasible* solution, but may not be the *optimal* solution to Problem 1. In general, the solution $(\hat{P}, \hat{A}, \hat{B}, C)$ is conservative for Problem 1. In the following subsection, we aim at finding a better approximation of Problem 1.

3.3 Sequential Convex Approximation of Problem 1

In this subsection, we give the efficient solution method for Problem 1 based on the overbounding method proposed by Sebe (2018). In the overbounding method, the inner approximations of nonlinear matrix inequalities are sequentially constructed. This sequential method contributes to gradually reduce the conservativeness of the solution of Problem 2.

Suppose that Problem 2 is feasible and that the feasible solution to Problem 1, denoted by $(\hat{P}, \hat{A}, \hat{B})$, is constructed. Then, we try to update the *initial guess* $(\hat{P}, \hat{A}, \hat{B})$ to reduce the conservativeness, i.e., to further reduce $J_1(A,B)$. First, we transform the decision variables of Problem 1, denoted by (P,A,B), into $(\Delta P, \Delta A, \Delta B)$ as follows.

$$P = \hat{P} + \Delta P, \quad A = \hat{A} + \Delta A, \quad B = \hat{B} + \Delta B.$$

Further, we let G and H be additional decision variables. With those G and H, we define the inequality condition described by

$$\mathrm{He}\left(\begin{bmatrix} Q(\Delta P, \Delta A, \Delta B) & \begin{bmatrix} 0 \\ \Delta P \end{bmatrix} & 0 \\ 0 & -G & G \\ -H\,[\,\Delta A\ \Delta B\ 0\,] & 0 & -H \end{bmatrix} \right) \prec 0, \quad (21)$$

where $Q(\Delta P, \Delta A, \Delta B)$ is given by (22) and $F(\Delta P)$ is

$$F(\Delta P) = \begin{bmatrix} \hat{P} + \Delta P - C^{\mathrm{T}}\Xi_{11}C & -C^{\mathrm{T}}\Xi_{12} \\ -\Xi_{12}^{\mathrm{T}}C & -\Xi_{22} \end{bmatrix}.$$

Table 1. Notation of optimal solutions

Problem	Solution
Unconstrained (8)	$(A^\dagger, B^\dagger, C^\dagger)$
Problem 1	(P^*, A^*, B^*, C^*)
Problem 2	$(\hat{P}, \hat{R}, \hat{S}) \to (\hat{P}, \hat{A}, \hat{B})$
Problem 3	$(\Delta\bar{P}, \Delta\bar{A}, \Delta\bar{B}) \to (\bar{P}, \bar{A}, \bar{B})$

We show that (21) is a sufficient condition for (15) as stated in the following proposition.

Proposition 2. Suppose (21) holds. Then, letting $(P,A,B) = (\hat{P}+\Delta P, \hat{A}+\Delta A, \hat{B}+\Delta B)$, it holds that (15).

The proof follows Proposition 2 in the work by Sebe (2018) and is omitted in this paper. Furthermore, it should be noted that (21) is linear in $(\Delta P, \Delta A, \Delta B, G)$. This implies that for any fixed H, (21) is in the form of LMIs and is numerically tractable.

Recall $J_1(A,B)$ of (9) to obtain the expression

$$J_1(\hat{A}+\Delta A, \hat{B}+\Delta B) = \left\| \Psi_{k+1} - \left[\hat{A}+\Delta A,\ \hat{B}+\Delta B \right] \begin{bmatrix} \Psi_k \\ U_k \end{bmatrix} \right\|_F^2. \quad (23)$$

Then, the problem of finding $(\Delta P, \Delta A, \Delta B, G)$ that minimizes $J_1(\hat{A}+\Delta A, \hat{B}+\Delta B)$ under the constraint (21) based on the initial guess $(\hat{P}, \hat{A}, \hat{B})$ is stated as follows

Problem 3. Given the system matrix C, the real symmetric matrix Ξ, the data-matrices $(U_k, \Psi_k, \Psi_{k+1})$, the feasible solution to Problem 1, denoted by $(\hat{P}, \hat{A}, \hat{B})$, and the real matrix H, solve the optimization problem:

$$\min_{\Delta P, \Delta A, \Delta B, G} \quad J_1(\hat{A}+\Delta A, \hat{B}+\Delta B)$$
$$\text{sub to} \quad (21),\ \hat{P}+\Delta P \succ 0.$$

With the optimal solution $(\Delta\bar{P}, \Delta\bar{A}, \Delta\bar{B})$ to Problem 3, we obtain the matrices

$$\bar{P} = \hat{P}+\Delta\bar{P}, \quad \bar{A} = \hat{A}+\Delta\bar{A}, \quad \bar{B} = \hat{B}+\Delta\bar{B}.$$

The notation on the optimal solutions to Problems 1–3 are summarized on Table 1.

Note that the solution $(\bar{P}, \bar{A}, \bar{B})$ is the less conservative solution to Problem 1 than the initial guess $(\hat{P}, \hat{A}, \hat{B})$ for any real matrix H satisfying

$$H + H^{\mathrm{T}} \succ 0. \quad (24)$$

This fact is mathematically stated in the following proposition.

Proposition 3. Suppose that Problem 2 is feasible. Then, for any real matrix H satisfying the condition (24), Problem 3 is feasible. In addition, if $(\hat{P}, \hat{A}, \hat{B})$ is not the solution to Problem 1, the strict inequality

$$J_1(\bar{A}, \bar{B}) < J_1(\hat{A}, \hat{B}) \quad (25)$$

holds.

The proof is omitted in this paper.

On the basis of the fact stated in Proposition 3, a sequential algorithm of solving Problem 1 is proposed. Suppose that Problem 3 with $(\hat{P}, \hat{A}, \hat{B}, H) = (\bar{P}_i, \bar{A}_i, \bar{B}_i, H_i)$ has the optimal solution $(\Delta\bar{P}_i, \Delta\bar{A}_i, \Delta\bar{B}_i, \bar{G}_i)$. Consider the updating law

$$(\bar{P}_{i+1}, \bar{A}_{i+1}, \bar{B}_{i+1}, H_{i+1}) \leftarrow (\bar{P}_i+\Delta\bar{P}_i, \bar{A}_i+\Delta\bar{A}_i, \bar{B}_i+\Delta\bar{B}_i, \bar{G}_i). \quad (26)$$

$$Q(\Delta P, \Delta A, \Delta B) = -\frac{1}{2}\begin{bmatrix} -F(\Delta P) & 0 \\ 0 & \hat{P}+\Delta P \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ -\hat{P}\hat{A} & -\hat{P}\hat{B} & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ -\Delta P\hat{A} & -\Delta P\hat{B} & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ -\hat{P}\Delta A & -\hat{P}\Delta B & 0 \end{bmatrix}. \quad (22)$$

Then, by sequentially solving Problem 3 with the updated $(\hat{P}, \hat{A}, \hat{B}, H) = (\bar{P}_{i+1}, \bar{A}_{i+1}, \bar{B}_{i+1}, H_{i+1})$, we obtain the solution $(\Delta\bar{P}_{i+1}, \Delta\bar{A}_{i+1}, \Delta\bar{B}_{i+1}, \bar{G}_{i+1})$, which generates the less conservative solution to Problem 1.

We propose to sequentially solve Problem 3 with updated $(\bar{P}_{i+1}, \bar{A}_{i+1}, \bar{B}_{i+1}, H_{i+1})$ to obtain the less conservative solution to Problem 1. The sequential solution method is summarized in Algorithm.

Algorithm

1. Find the solution $(\hat{P}, \hat{R}, \hat{S}) = (\hat{P}_0, \hat{R}_0, \hat{S}_0)$ to Problem 2 and obtain $\bar{A}_0 = \hat{P}_0^{-1}\hat{R}_0$ and $\bar{B}_0 = \hat{P}_0^{-1}\hat{S}_0$. In addition, let $H_0 = I$ and $i = 0$.
2. Given $(\hat{P}, \hat{A}, \hat{B}, H) = (\bar{P}_i, \bar{A}_i, \bar{B}_i, H_i)$, find the solution $(\Delta\bar{P}_i, \Delta\bar{A}_i, \Delta\bar{B}_i, \bar{G}_i)$ to Problem 3.
3. Apply (26) to obtain $(\bar{P}_{i+1}, \bar{A}_{i+1}, \bar{B}_{i+1}, H_{i+1})$.
4. If $|J_1(\bar{A}_{i+1}, \bar{B}_{i+1}) - J_1(\bar{A}_i, \bar{B}_i)| < \epsilon$ for a positive constant ϵ, then terminate the algorithm.
5. Set $i \leftarrow i+1$ and go to 2.

---

4. NUMERICAL EXPERIMENT

In this section, we demonstrate the procedure of learning a nonlinear dynamical system by applying the proposed algorithm. Consider a continuous-time nonlinear dynamical system described by

$$\begin{cases} \dot{x}_1(t) = x_2(t), \\ \dot{x}_2(t) = -2x_2(t) + x_1(t)\cos(x_1(t)+x_2(t)) + u(t), \\ y(t) = x_2(t). \end{cases} \quad (27)$$

It is known that the system is dissipative with respect to $(\Xi_{11}, \Xi_{12}, \Xi_{22}) = (0, -1, 0)$, i.e., the system is passive (see e.g., the work by Zakeri and Antsaklis (2019)). In this experiment, we aim at accurately learning the dynamical system in the Koopman model (7), while incorporating the passivity property.

In the experimental setup, we consider that the time series of $x(t)$, $u(t)$, and $y(t)$ are sampled at each 0.01 time interval from the system (27), which are denoted by $\{x(k)\}$, $\{u(k)\}$, and $\{y(k)\}$, respectively. The input series $\{u(k)\}$ for learning are determined from randomly selected values from the uniform distribution in $[-1, 1]$. Then, the state and output series $\{x(k)\}$ and $\{y(k)\}$ are measured synchronously. In total, the data at 5000 samples are obtained.

We try to apply Algorithm to the data $\{x(k)\}$, $\{u(k)\}$, and $\{y(k)\}$. To this end, first, we let $(\Xi_{11}, \Xi_{12}, \Xi_{22}) = (0, -1, -0.2)$ and its corresponding dissipativity constraint be defined to inherit the *a priori* information on the passivity in a relaxed form. Furthermore, let the lifting function $\psi(x(k))$ be composed of the state $x(k) = [x_1(k), x_2(k)]^{\mathrm{T}}$ and thin plate spline radial basis functions $\psi_i(x(k))$, $i \in \{1, 2, \ldots, 8\}$, where $\psi_i(x(k))$ is given by

$$\psi_i(x(k)) = \|x(k) - r_i\|_2^2 \ln \|x(k) - r_i\|_2$$

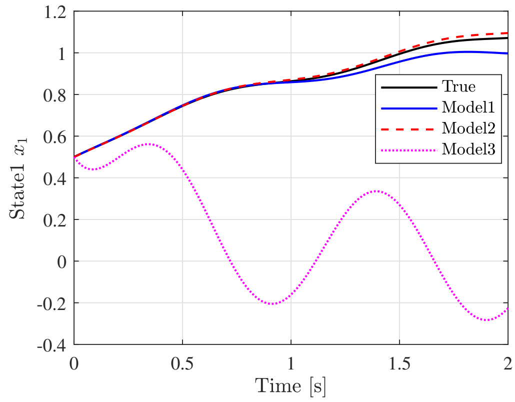

Fig. 1. State trajectory $x_1(k)$.

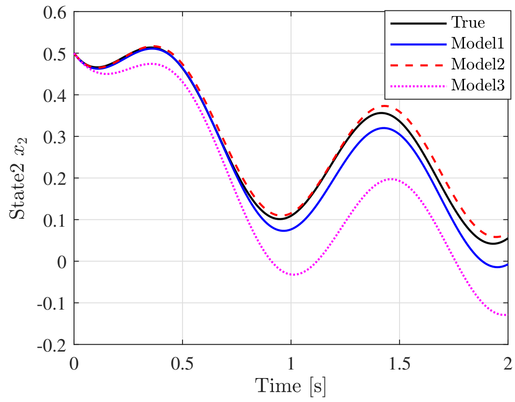

Fig. 2. State trajectory $x_2(k)$.

and the values of r_i are selected randomly from the uniform distribution on the unit box. Then, the lifting function $\psi(x(k))$ is described by

$$\psi(x(k)) = [\,x(k)\ \psi_1(x(k))\ \cdots\ \psi_8(x(k))\,]^{\mathrm{T}} \in \mathbb{R}^{2+8}. \quad (28)$$

By applying Algorithm, we constructed the dissipativity-constrained Koopman model, which approximates the nonlinear system (27) and is called Model 1. In addition, we constructed two different models by using the same time series data: (Model 2) one is the no-constrained Koopman model, which is simply constructed by solving (8), while (Model 3) the other is the dissipativity-constrained *linear* model, which is based on $\psi(x(k)) = x(k)$ and is constructed by applying the learning method proposed by Abe et al. (2016).

First, to show the model accuracy, the three models are compared with the true nonlinear system (27). The result of the frequency response against the sin-wave input is illustrated in Figs. 1 and 2, where the state trajectory of the models is shown. In the figures, the black solid, blue solid, red dashed, and pink dotted lines represent the state of the true system, Model 1 (proposed model), Model 2, and Model 3, respectively. We see from Fig. 1 that Models 1 and 2, i.e., the Koopman models, accurately express the nonlinear behavior generated by (27), while Model 3, i.e., the linear model, is not. The lifting function with the basis (28) contributes to improving the ability of model expression.

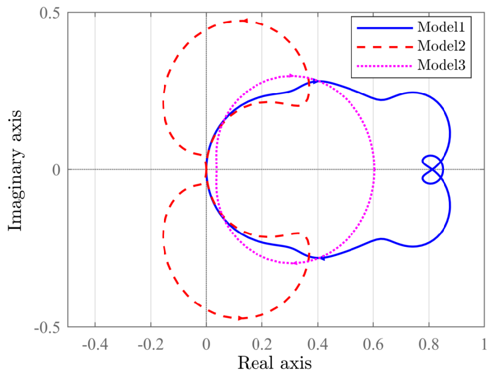

Fig. 3. Comparison of Nyquist plot

Next, to show the validity of the dissipativity constraint, we define the transfer function for the Koopman model. Letting $G(e^{j\omega T}) = C(e^{j\omega T}I - A)^{-1}B$, we reduce the dissipativity constraint, characterized by $(\Xi_{11},\Xi_{12},\Xi_{22})=(0,-1,-0.2)$, to

$$\mathrm{Re}[G(e^{j\omega T})] \geq -0.1, \quad \forall \omega \in \mathbb{R}.$$

Then, the Nyquist plot of the $G(e^{j\omega T})$ for the three models is illustrated in Fig. 3. As illustrated in the figure, Models 1 and 3 satisfy the dissipativity constraint, while Model 2 violates. This shows that the dissipativity constraint imposed on the learning problems is valid.

5. CONCLUSION

This paper addressed the learning problem of nonlinear dynamical systems with incorporating the *a priori* information on the quadratic dissipativity. The problem was reduced to the data-driven approximation of the Koopman operator under the dissipativity constraint, which was called Problem 1 in this paper. Then, the solution method to the problem was given and summarized in Algorithm. There are two main contributions of this paper. 1) One is in this numerically efficient algorithm, which sequentially solves LMIs. 2) The other is the performance analysis of the algorithm, which is stated in Proposition 3. In the analysis, it is guaranteed that the Koopman model constructed by the algorithm fits given data more accurately than the model defined at any initial guess.

ACKNOWLEDGEMENTS

This work was supported by Grant-in-Aid for Young Scientists (B), No. 17K14704 from JSPS.

REFERENCES

Abe, Y., Inoue, M., and Adachi, S. (2016). Subspace identification method incorporated with a priori information characterized in frequency domain. In *2016 European Control Conference (ECC)*, 1377–1382. IEEE.

Alenany, A., Shang, H., Soliman, M., and Ziedan, I. (2011). Improved subspace identification with prior information using constrained least squares. *IET Control Theory & Applications*, 5(13), 1568–1576.

Brogliato, B., Lozano, R., Maschke, B., and Egeland, O. (2007). Dissipative systems analysis and control. *Theory and Applications*, 2.

De la Rosa, E. and Yu, W. (2016). Randomized algorithms for nonlinear system identification with deep learning modification. *Information Sciences*, 364, 197–212.

Goethals, I., Van Gestel, T., Suykens, J., Van Dooren, P., and De Moor, B. (2003). Identification of positive real models in subspace identification by using regularization. *IEEE Transactions on Automatic Control*, 48(10), 1843–1847.

Hill, D. and Moylan, P. (1976). The stability of nonlinear dissipative systems. *IEEE Transactions on Automatic Control*, 21(5), 708–711.

Hill, D.J. and Moylan, P.J. (1977). Stability results for nonlinear feedback systems. *Automatica*, 13(4), 377–382.

Hoagg, J.B., Lacy, S.L., Erwin, R.S., and Bernstein, D.S. (2004). First-order-hold sampling of positive real systems and subspace identification of positive real models. In *Proceedings of the 2004 American control conference*, volume 1, 861–866. IEEE.

Inoue, M. (2019). Subspace identification with moment matching. *Automatica*, 99, 22–32.

Jin, X., Shao, J., Zhang, X., An, W., and Malekian, R. (2016). Modeling of nonlinear system based on deep learning framework. *Nonlinear Dynamics*, 84(3), 1327–1340.

Kevrekidis, I., Rowley, C., and Williams, M. (2015). A kernel-based method for data-driven Koopman spectral analysis. *Journal of Computational Dynamics*, 2(2), 247–265.

Koopman, B.O. (1931). Hamiltonian systems and transformation in hilbert space. *Proceedings of the National Academy of Sciences of the United States of America*, 17(5), 315–318.

Korda, M. and Mezić, I. (2018). Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control. *Automatica*, 93, 149–160.

Lacy, S.L. and Bernstein, D.S. (2003). Subspace identification with guaranteed stability using constrained optimization. *IEEE Transactions on automatic control*, 48(7), 1259–1263.

Miller, D.N. and De Callafon, R.A. (2013). Subspace identification with eigenvalue constraints. *Automatica*, 49(8), 2468–2473.

Okada, M. and Sugie, T. (1996). Subspace system identification considering both noise attenuation and use of prior knowledge. In *Proceedings of 35th IEEE Conference on Decision and Control*, volume 4, 3662–3667. IEEE.

Sebe, N. (2018). Sequential convex overbounding approximation method for bilinear matrix inequality problems. *IFAC-PapersOnLine*, 51(25), 102–109.

Willems, J.C. (1972). Dissipative dynamical systems part i: General theory. *Archive for Rational Mechanics and Analysis*, 45(5), 321–351. doi:10.1007/BF00276493.

Williams, M.O., Kevrekidis, I.G., and Rowley, C.W. (2015). A data–driven approximation of the Koopman operator: Extending dynamic mode decomposition. *Journal of Nonlinear Science*, 25(6), 1307–1346.

Yoshimura, S., Matsubayashi, A., and Inoue, M. (2019). System identification method inheriting steady-state characteristics of existing model. *International Journal of Control*, 92(11), 2701–2711.

Zakeri, H. and Antsaklis, P.J. (2019). Passivity and passivity indices of nonlinear systems under operational limitations using approximations. *International Journal of Control*, (just-accepted), 1–20.