

Learning Deep Dissipative Dynamics

Yuji Okamoto¹*, Ryosuke Kojima^{1,2}*

¹ Kyoto University, Japan

² RIKEN BDR, Japan

{okamoto.yuji.2c, kojima.ryosuke.8e}@kyoto-u.ac.jp

Abstract

This study challenges strictly guaranteeing “dissipativity” of a dynamical system represented by neural networks learned from given time-series data. Dissipativity is a crucial indicator for dynamical systems that generalizes stability and input-output stability, known to be valid across various systems including robotics, biological systems, and molecular dynamics. By analytically proving the general solution to the non-linear Kalman–Yakubovich–Popov (KYP) lemma, which is the necessary and sufficient condition for dissipativity, we propose a differentiable projection that transforms any dynamics represented by neural networks into dissipative ones and a learning method for the transformed dynamics. Utilizing the generality of dissipativity, our method strictly guarantee stability, input-output stability, and energy conservation of trained dynamical systems. Finally, we demonstrate the robustness of our method against out-of-domain input through applications to robotic arms and fluid dynamics.

Code —

<https://github.com/kojima-r/DeepDissipativeModel>

Extended version —

<https://doi.org/10.48550/arXiv.2408.11479>

1 Introduction

Dissipativity extends the concept of Lyapunov stability to input-output dynamical systems by considering “energy” (Brogliato et al. 2020). In input-output systems, the relationship between the externally supplied energy and the dissipated energy plays an important role. The theory of dissipativity has wide applications, including electrical circuits (Ortega and Ortega 1998), mechanical systems (Hatanaka et al. 2015), and biological systems (Goldbeter 2018). Considering the inflow, outflow, and storage of energy in a system provides crucial insights for applications such as stability analysis, controller design, and complex interconnected systems.

Theoretically, dissipativity in input-output dynamical systems is defined by the time evolution of inputs $u(t)$, outputs

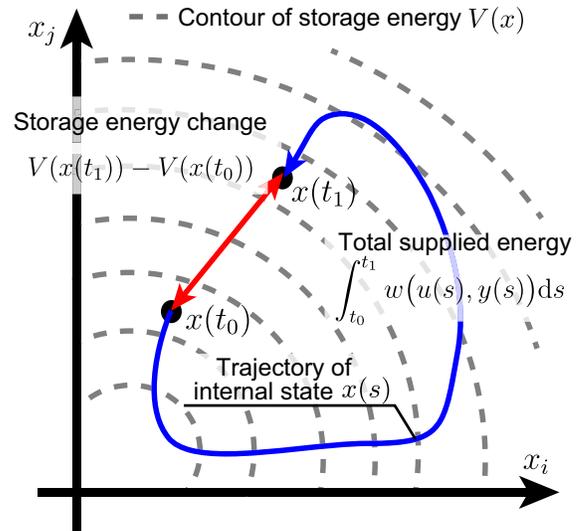


Figure 1: Sketch of the dissipativity: The red difference in storage energy is less than the total energy supplied along the blue line, which represents the trajectory of the internal state $x(s)$.

$y(t)$, and internal states $x(t)$. The input-output system is dissipative if the following inequality is satisfied:

$$\underbrace{V(x(t_1)) - V(x(t_0))}_{\text{Storage Energy Changes}} \leq \underbrace{\int_{t_0}^{t_1} w(u(s), y(s)) ds}_{\text{Total Energy Supplies}}, \quad (1)$$

where $V(x(t))$ is called the storage energy and $w(u(t), y(t))$ is called the supply rate. The left side represents the change in storage energy from the initial state $x(t_0)$ to the final state $x(t_1)$, while the right side signifies the total supplied energy from t_0 to t_1 (See Figure 1). In Newtonian mechanics, if $V(x)$ is defined as the mechanical energy and $w(u, y)$ is defined as the product of external force u and velocity y , i.e., $w(u, y) = uy$, this case corresponds to the principle of energy conservation (Stramigioli 2006). In this context, the integral of the supply rate represents the “work” done by the external force u .

In this study, we propose an innovative method for learning dynamical systems described by neural networks from

*These authors contributed equally.

time-series data when the system is known a priori to possess dissipativity. Considering the entire space of the dynamical system described by neural networks, we introduce a transformation of the system by a projection map onto the subspace satisfying dissipativity. We emphasize that this projection can be applied to the dynamical systems consisting of any differentiable neural networks. By incorporating this projection into the gradient-based optimization of neural networks, our method allows fitting dissipative dynamics to time-series data.

By configuring the supply rate $w(\cdot, \cdot)$ in the dissipativity, users can design models integrating well-established prior knowledge such as properties of dynamical systems or information from physical systems. According to the properties of the target dynamical system, such as internal stability, input-output stability, and energy conservation, the supply rate $w(\cdot, \cdot)$ can be constrained. Within physical systems, the supply rate $w(\cdot, \cdot)$ can be derived from the principle of energy conservation.

Real-world environments often present inputs that vary from the input-output datasets used during model training, for example due to dataset shifts. Our proposed method guarantees that the trained model strictly satisfies dissipativity for any input time-series data, thereby maintaining robust performance on out-of-domain input. In this study, we verified the effectiveness of our method, particularly its robustness to out-of-domain input, using both linear and nonlinear systems, including an n -link pendulum (a principal model of robotic arms) and the behavior of viscous fluid around a cylinder.

The contributions of this study are as follows:

- (i) We analytically derived a general solution to the nonlinear KYP lemma and a differentiable projection from the dynamical systems represented by neural networks to a subspace of dissipative ones.
- (ii) We proposed a learning method for strictly dissipative dynamical systems using the above projection.
- (iii) We showed that our learning method generalizes existing methods that guarantee internal stability and input-output stability.
- (iv) We confirmed the effectiveness of our method with three experiments with benchmark data.

2 Related Work

Learning Stable Dynamics. In recent years, numerous methods have been proposed for learning models with a priori properties, such as system stability, rather than relying solely on data (Blocher, Saveriano, and Lee 2017; Khansari-Zadeh and Billard 2011; Umlauf and Hirche 2017). With the advent of deep learning, techniques have been developed to enhance the stability of loss functions compatible with gradient-based learning (Richards, Berkenkamp, and Krause 2018).

Manek et al. tackled the same internal system but introduced a novel method that guarantees the stability without depending on loss optimization by analytically guaranteeing internal stability (Manek and Zico Kolter 2019). This approach was further extended to apply positive invariant sets,

such as limit cycles and line attractors, to ensure internal stability (Takeishi and Kawahara 2021). Additionally, this analytical approach has been developed for closed-loop systems, ensuring their stability through an SMT solver (Chang, Roohi, and Gao 2019).

Lawrence et al. utilized stochastic dynamical systems, emphasizing internal stability and maintaining it through a loss-based approach (Lawrence et al. 2020). Similarly, another method has been proposed for state-space models, focusing on input-output stability and ensuring this through projection (Kojima and Okamoto 2022).

Unlike the above approaches, techniques that imposes constraints on the architecture of neural networks to guarantee energy dissipativity has also been proposed (Xu and Sivarajani 2023; Sosanya and Greydanus 2022).

Hamiltonian NN. Related to the learning of stable systems, Hamiltonian Neural Networks (HNNs) incorporate the principle of energy conservation into their models (Greydanus, Dzamba, and Yosinski 2019). Hamiltonian dynamical systems maintain conserved energy, allowing HNNs to learn Hamiltonian functions to predict time evolution. This method ensures that the model adheres to the conservation of energy law, resulting in physically accurate predictions.

Conversely, some systems exhibit decreasing energy over time without external input, a characteristic known as “dissipation.” This property is prevalent in many real-world systems, particularly those involving thermodynamics and friction. Consequently, methods for learning systems with dissipation from data are gaining interest (Drgoňa et al. 2022).

By generalizing dissipation from energy to a broader positive definite function V , it can represent a unified concept encompassing input-output stability and Lyapunov stability. In this study, we adopted this broader interpretation of dissipativity, allowing us to understand the learning of systems that ensure stability-related properties in a unified framework. Hereafter, in this paper, we will use the term “dissipativity” without distinguishing between “dissipation” and “dissipativity.”

Neural ODE. The state-space dynamic system can be regarded as a differential equation, and our implementation actually uses neural ODE as an internal solver (Chen et al. 2018; Chen 2019). These techniques have been improved in recent years, including discretization errors and computational complexity. Although we used an Euler method for simplicity, we can expect that learning efficiency would be further improved by using these techniques. In this field, methods have been proposed that mainly learn various types of continuous-time dynamics from data. For example, extended methods for learning stochastic dynamics have been proposed (Kidger et al. 2020; Morrill et al. 2021)

3 Background

This study deals with continuous-time state-space models as input-output systems using a nonlinear Lipschitz continuous mapping $f(x) \in \mathbb{R}^n$ with $f(0) = 0$, continuous mappings $g(x) \in \mathbb{R}^{n \times m}$, $h(x) \in \mathbb{R}^l$ with $h(0) = 0$ and $j(x) \in \mathbb{R}^{l \times m}$

formed by neural networks:

$$\begin{aligned}\dot{x} &= f(x) + g(x)u, & x(0) &= x_0 \\ y &= h(x) + j(x)u\end{aligned}\quad (2)$$

where the internal state x , the input u , and the output y belong to a signal spaces that maps from time interval $[0, \infty)$ to the n , m , and l dimensional Euclidean space, respectively. Here, a tuple (f, g, h, j) is called a dynamics of the input-output system (2). Dissipativity is defined by the supply of energy through the input-output signals u, y and the change in storage energy depending on the internal state x .

Definition 1 (Dissipativity). *Considering a supply rate $w : \mathbb{R}^m \times \mathbb{R}^l \rightarrow \mathbb{R}$, there exist a differentiable positive semi-definite storage function $V : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$ such that the input-output system (2) satisfies the dissipative condition (1), then the system is dissipative.*

Due to the flexible definition of the supply rate $w(u, y)$, dissipativity can be precisely designed to match the energy conservation law of physical systems, as well as adapt to the properties of dynamical systems, such as internal stability.

For example in Newtonian mechanics, the sum of kinetic and potential energy can be regarded as the storage function $V(x)$. The supply rate $w(u, y)$ can be determined by the difference between the ‘‘work’’ done by external forces and energy dissipativity caused by air resistance or friction. This work is represented as the integral of the product of velocity and external force. This energy dissipativity due to friction can be expressed as a quadratic form of velocity. This supply rate belongs to a quadratic form of external input and observed velocity (see Appendix K).

Additionally, dissipativity is defined as an extension of internal stability and input-output stability. For $w(u, y) = 0$, it corresponds to internal stability (assuming V is positive definite), for $w(u, y) = \gamma^2 \|u\|^2 - \|y\|^2$, it corresponds to input-output stability ($\gamma > 0$ is the gain of input-output signals).

In general, the supply rate $w(u, y)$, as described in the above two paragraphs, is represented as a quadratic form of the input and output:

$$w(u, y) \triangleq [y^T, u^T] \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} y \\ u \end{bmatrix}\quad (3)$$

The supply rate parameters Q, S, R can be designed in a manner that corresponds to the energy conservation laws of physical systems and the properties of dynamical systems such as internal stability. When the supply rate can be expressed in this form, there exists a necessary and sufficient condition of dissipative dynamical systems, formulated as the following matrix equation:

Proposition 2 (Nonlinear KYP lemma (Brogliato et al. 2020, Theorem 4.101)). *Consider the input-output system (2) is reachable. The system (2) is dissipative if and only if there exists $\ell : \mathbb{R}^n \rightarrow \mathbb{R}^q$, $W : \mathbb{R}^n \rightarrow \mathbb{R}^{q \times m}$ and a differentiable positive semi-definite function $V : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$ such*

that

$$\begin{aligned}\nabla V^T(x)f(x) &= h^T(x)Qh(x) - \ell^T(x)\ell(x), \\ \frac{1}{2}\nabla V^T(x)g(x) &= h^T(x)(S + Qj(x)) - \ell^T(x)W(x), \\ W^T(x)W(x) &= R + j^T(x)S + S^Tj(x) + j^T(x)Qj(x), \\ &\forall x \in \mathbb{R}^n,\end{aligned}\quad (4)$$

where the nonlinear dynamical system is reachable if and only if for any x^* there exists $T \geq 0$ and u such that $x(0) = 0$ and $x(T) = x^*$.

Proof. Appendix A. \square

The maps ℓ and W represent the residuals of the time derivative of the storage function V and the supply rate $w(u, y)$ in the definition of dissipativity. The maps ℓ corresponds to terms independent of the input u , while W corresponds to terms linearly dependent on u (See detail in Appendix A).

Note that the maps ℓ, W and V satisfying the condition (4) is not unique. See Appendix B for a more detailed discussion about the freedom degree of the dissipativity condition.

Assuming $j \equiv 0$, the conditions for input-output stability can be easily derived from the nonlinear KYP lemma and is known as the Hamilton-Jacobi inequality (Details are provided in the Appendix C). Various dynamical systems except for the field of electronic circuits often lack a direct path j from input to output.

The nonlinear KYP lemma means the existence of a (non-unique) mapping from dissipative dynamics (f, g, h, j) to conditions-satisfying maps $\ell(x)$, $W(x)$, and $V(x)$. On the contrary, it has not been demonstrated whether there is a mapping from (ℓ, W, V) to the dynamics (f, g, h, j) satisfying dissipativity. If it is possible to derive a mapping from (ℓ, W, V) to dissipative dynamics (f, g, h, j) , then by managing (ℓ, W, V) , indirect constraining dissipative dynamics (f, g, h, j) could become possible.

4 Method

4.1 Projection-based Optimization Framework

The aim of this study is to learn strictly dissipative dynamics fitted to input-output data. We consider the subspace of all dissipative dynamics within the function space consisting of tuples of four nonlinear maps (f, g, h, j) constructed by neural networks. By projecting the dynamics (f, g, h, j) onto the subspace of dissipative dynamics, the resulting neural network-based dynamical system will inherently satisfy dissipativity. Consequently, by training the projected neural networks to fit the input-output data, both fitting accuracy and strict dissipativity are achieved. Considering a parametric subspace of dissipative dynamics, we introduce the parameterized projection onto this subspace.

Definition 3 (Dissipative projection). *Let \mathbb{S} be a function space of (f, g, h, j) , $\mathbb{S}_d \subset \mathbb{S}$ be the subspace satisfying dissipativity, and Θ be a parameter set. If the differentiable functional $\mathcal{P}_\theta : \mathbb{S} \rightarrow \mathbb{S}_\theta \subset \mathbb{S}_d$ satisfies*

$$\mathcal{P}_\theta \circ \mathcal{P}_\theta = \mathcal{P}_\theta, \quad \mathbb{S}_d = \bigcup_{\theta \in \Theta} \mathbb{S}_\theta, \quad (5)$$

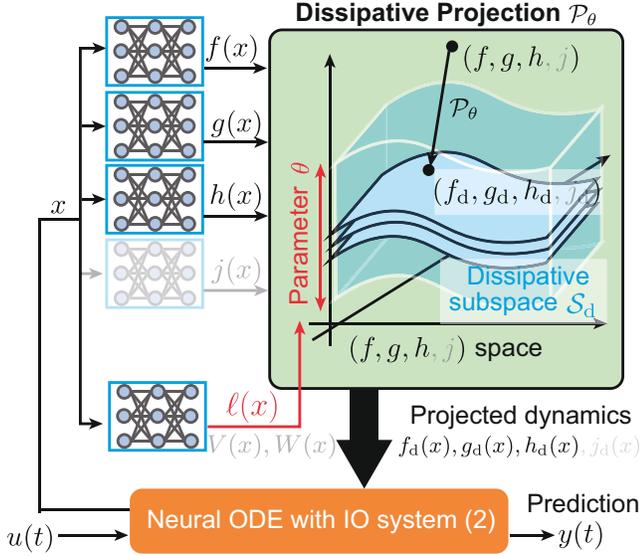


Figure 2: Sketch of the proposed method: The dynamics of the input-output system (f, g, h, j) is projected into a space with guaranteed dissipativity using dissipative projection \mathcal{P}_θ , and the output signal $y(t)$ is predicted using the projected dynamics (f_d, g_d, h_d, j_d) and the input signal $u(t)$.

then, \mathcal{P}_θ is called a dissipative projection.

The nonlinear KYP lemma serves as a necessary and sufficient condition for dissipativity, meaning that \mathbb{S}_d aligns with the entirety of dynamics satisfying this condition. By fixing the maps (ℓ, W, V) , we determine a subspace of dynamics (f, g, h, j) that complies with the equation (4) of the nonlinear KYP lemma. By unfixing the maps (ℓ, W, V) , the union of all of such subspaces related to (ℓ, W, V) corresponds to the entire set of dissipative dynamics. So, the maps (ℓ, W, V) can be regarded as the parameter θ on Definition 3 (See Figure 2).

By jointly learning the pre-projected dynamics (f, g, h, j) and the indirect parameter θ , it becomes possible to optimize dissipative dynamics. The formulation for learning strictly dissipative dynamics is established using the dissipative projection \mathcal{P}_θ as follows:

Problem 4. Let $D \triangleq \{u_i, y_i^*\}_{i=1}^N$ be a dataset and \mathcal{P}_θ be a dissipative projection. Our problem is written as

$$\underset{(f, g, h, j) \in \mathbb{S}, \theta \in \Theta}{\text{minimize}} \quad \mathbb{E}_{(u, y^*) \in D} [\|y^* - y\|^2] \quad (6)$$

where y is the prediction result by the input signal u and the input-output system (2) from projected dynamics

$$(f_d, g_d, h_d, j_d) \triangleq \mathcal{P}_\theta(f, g, h, j).$$

In the following sections, we analytically derive concrete dissipative projections \mathcal{P}_θ based on the nonlinear KYP lemma. To derive the dissipative projection with parameters (ℓ, W, V) , we generally solve the equation (4) in the nonlinear KYP lemma. Therefore, in the next section, we derive

the general solution to the matrix equation of the nonlinear KYP lemma, and in Section 4.3, we derive a dissipative projection using the parameters (ℓ, W, V) based on this general solution. Finally, we introduce a loss function to realize efficient learning for dissipative input-output system (2).

4.2 General Solution of Nonlinear KYP Lemma

For any maps (ℓ, W, V) , equations (4) in the nonlinear KYP lemma are written as a quadratic matrix equation (QME) form of the dynamics (f, g, h, j) :

$$X^T A X + B^T X + X^T B + C = 0 \quad (7)$$

where

$$X \triangleq \begin{bmatrix} f & g \\ h & j \end{bmatrix}, \quad A \triangleq \begin{bmatrix} 0 & 0 \\ 0 & -Q \end{bmatrix}, \quad (8)$$

$$B \triangleq \begin{bmatrix} \frac{1}{2} \nabla V & 0 \\ 0 & -S \end{bmatrix}, \quad C \triangleq \begin{bmatrix} \|\ell\|^2 & \ell^T W \\ W^T \ell & W^T W - R \end{bmatrix}.$$

The general solution of this QME presents the following.

Lemma 5. Assuming Q is a negative definite matrix, if $R - S^T Q^{-1} S - W^T W$ is a positive semi-definite matrix, then the QME (7) exists a solution, and the general solution is written as

$$f = P_{\nabla V}^c \hat{f} + \frac{\nabla V}{\|\nabla V\|^2} (\hat{h}^T Q \hat{h} - \|\ell\|^2) \quad (9a)$$

$$g = P_{\nabla V}^c \hat{g} + 2 \frac{\nabla V}{\|\nabla V\|^2} (\hat{h}^T (S + Q \hat{j}) - \ell^T W) \quad (9b)$$

$$h = \hat{h}, \quad j = \hat{j} \quad (9c)$$

where $P_{\nabla V}^c$ is the projection onto the complementary of subspace spanned by the vector ∇V which defined as

$$P_{\nabla V}^c \triangleq I_n - \frac{1}{\|\nabla V\|^2} \nabla V \nabla V^T.$$

The intermediate variables $(\hat{f}, \hat{g}, \hat{h}, \hat{j})$ are parameters in the solution of this QME (7), such that \hat{j} satisfies the following ellipsoidal condition:

$$\begin{aligned} (\hat{j} + Q^{-1} S)^T (-Q) (\hat{j} + Q^{-1} S) \\ = R - S^T Q^{-1} S - W^T W. \end{aligned} \quad (10)$$

Proof. See Appendix D. \square

The solution of the QME is divided into the null space and non-null space of the matrix A on the equation (8). Since $\begin{bmatrix} f \\ g \\ 0 \\ 0 \end{bmatrix}$ corresponds to the null space of the matrix A , it is the solution of the linear equation and $P_{\nabla V}^c[\hat{f}, \hat{g}]$ is the complementary space. Considering the non-null space of A , j is a matrix on the ellipsoidal sphere by a positive definite matrix $-Q$, and the existence condition is that the radius of this ellipsoidal $R - S^T Q^{-1} S - W^T W$ is a positive semi-definite matrix.

This lemma assumes that Q is a negative definite matrix, but it can similarly be shown when Q is a positive definite matrix. If Q has eigenvalues with both positive and negative

signs, or if a complementary space exists, the solution needs to be written for each eigenspace and becomes complicated.

The result of Lemma 5 implies that the entire set of dissipative dynamics \mathbb{S}_d is determined by the parameters (ℓ, W, V) and the intermediate variable $(\hat{f}, \hat{g}, \hat{h}, \hat{j})$ in the general solution. Noting that W can be reduced as a map of \hat{j} derived from the ellipsoidal condition (10), the entire set of dissipative dynamics \mathbb{S}_d is partitioned by only two parameters (ℓ, V) .

In the next section, we explicitly derive the differentiable projection onto the parametric subspace $\mathbb{S}_{\ell, V}$ of dissipative dynamics, excluding the direct path from input to output.

4.3 Projection onto Dissipative Dynamics Subspace

Based on Lemma 5, this section derives the projection onto the subspace of dissipative dynamics $\mathbb{S}_{\ell, V}$ for any given mappings ℓ and V . In many applications, the direct path from input to output j is often excluded ($j \equiv 0$). In such cases, the negative definite matrix Q assumed in Lemma 5 is no longer required. The following theorem proposes a projection of dissipative dynamics assuming $j \equiv 0$.

If the direct path j is not excluded, it is necessary to construct a projection that satisfies the ellipsoidal condition (10) for j . The general case of j , the projection onto the subspace of dissipative dynamics $\mathbb{S}_{\ell, V}$ is shown in Appendix E.

Theorem 6 (Dissipative Projection). *Assume that R is positive semi-definite matrices. The following map $\mathcal{P}_{\ell, V} : (f, g, h) \mapsto (f_d, g_d, h_d)$ satisfying*

$$f_d = P_{\nabla V}^c f + \frac{h^T Q h - \|\ell\|^2}{\|\nabla V\|^2} \nabla V, \quad (11)$$

$$g_d = P_{\nabla V}^c g + \frac{2}{\|\nabla V\|^2} \nabla V (h^T S - \ell^T \sqrt{R}), \quad (12)$$

$$h_d = h \quad (13)$$

is a dissipative projection, where $\ell : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $V : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$.

Proof. See Appendix F. \square

Note that \sqrt{R} is the root of a positive semi-definite matrix R , satisfying $\sqrt{R}\sqrt{R} = R$ and \sqrt{R} is symmetric.

Projections that strictly guarantee internal stability, input-output stability, and energy conservation are achieved by designing the supply rate parameters (Q, S, R) . The projection that guarantees internal stability coincide with the literature (Manek and Zico Kolter 2019), and the projection that guarantees input-output stability corresponds with another study (Kojima and Okamoto 2022). For details on the differences from previous studies, please refer to the Appendix J. Below, we demonstrate differentiable projections that guarantee each of these time-series characteristics.

Corollary 7 (Stable Projection). *The following map $\mathcal{P}_V : (f, g, h) \mapsto (f_d, g_d, h_d)$:*

$$\begin{aligned} f_d &= f - \frac{\nabla V}{\|\nabla V\|^2} \text{ReLU}(\nabla V^T f), \\ g_d &= g, \quad h_d = h \end{aligned} \quad (14)$$

is a projection into stable dynamics, where V is positive definite.

Proof. When $Q = R = S = 0$, it is derived from the theorem. \square

Corollary 8 (Input-Output Stable Projection). *The following map $\mathcal{P}_{\ell, V} : (f, g, h) \mapsto (f_d, g_d, h_d)$:*

$$\begin{aligned} f_d &= P_{\nabla V}^c f - \frac{\nabla V}{\|\nabla V\|^2} (\|h\|^2 + \|\ell\|^2), \\ g_d &= P_{\nabla V}^c g - 2\gamma \frac{\nabla V \ell^T}{\|\nabla V\|^2} \\ h_d &= h \end{aligned} \quad (15)$$

is a projection into input-output (\mathcal{L}_2) stable dynamics and the $\gamma > 0$ is the input-output gain.

Proof. When $Q = -I_l$, $S = 0$, and $R = \gamma^2 I_m$, it is derived from the theorem. \square

The definition of dissipativity is expressed as an inequality involving the integral of the supply rate and the change of the storage function. Assuming $R = 0$ and $\ell \equiv 0$, this becomes an equality condition. This allows for the construction of a projection that strictly preserves the energy conservation law.

Corollary 9 (Energy Conservation Projection). *Assuming $R = 0$, if the following mapping $\mathcal{P}_V : (f, g, h) \mapsto (f_d, g_d, h_d)$ is given by*

$$\begin{aligned} f_d &= P_{\nabla V}^c f + \frac{\nabla V}{\|\nabla V\|^2} h^T Q h, \\ g_d &= P_{\nabla V}^c g + 2 \frac{\nabla V}{\|\nabla V\|^2} h^T S, \\ h_d &= h \end{aligned} \quad (16)$$

then the input-output system (2) satisfies

$$V(x(t_1)) - V(x(t_0)) = \int_{t_0}^{t_1} w(u(s), y(s)) ds.$$

Proof. See Appendix G. \square

The energy conservation projection supports the Hamiltonian equations, which conserve energy, and the port-Hamiltonian systems, where energy exchange is explicitly defined. In this context, the storage function V corresponds to the Hamiltonian, and the supply rate $w(u, y)$ corresponds to the energy dissipation in the port-Hamiltonian system.

A similar concept to dissipativity in evaluating input-output systems is passivity. Since passivity can be described as the exchange of energy, it can naturally be addressed in this study by adjusting the supply rate parameter for dissipativity (see Appendix H)

Note that dissipative projections are not unique because they depend on space metrics. Additionally, since the dissipative constraint is nonlinear, explicitly describing the underlying space metric is difficult. For instance, existing study (Kojima and Okamoto 2022) presents projections onto a non-Hilbert metric spaces under a simple quadratic constraint called input-output stability. For further details, please refer to the Appendix I.

Train	Test	Naive	Stable	IO stable	Conservation	Dissipative
Rectangle (N=100)	Rectangle	0.250 ± 0.184	0.252 ± 0.181	0.194 ± 0.095	0.077 ± 0.066	0.212 ± 0.144
	Step	0.205 ± 0.195	0.240 ± 0.203	0.225 ± 0.115	0.046 ± 0.021	0.197 ± 0.147
	Random	0.049 ± 0.044	0.047 ± 0.037	0.068 ± 0.036	0.023 ± 0.031	0.040 ± 0.028
Rectangle (N=1000)	Rectangle	0.029 ± 0.000	0.029 ± 0.000	0.029 ± 0.000	0.029 ± 0.000	0.060 ± 0.061
	Step	0.024 ± 0.000	0.024 ± 0.001	0.024 ± 0.001	0.024 ± 0.001	0.039 ± 0.029
	Random	0.005 ± 0.001	0.009 ± 0.003	0.007 ± 0.002	0.006 ± 0.003	0.021 ± 0.030

Table 1: The prediction error (RMSE) of the mass-spring-damper benchmark

4.4 Loss Function

The optimization problem (6) based on the dissipative projection requires careful attention to learning methods, as there is a degree of freedom in the internal parameters. Here, we define the regularized loss function as follows:

$$\text{Loss} \triangleq \mathbb{E}_{\mathcal{D}} \left[\|y^* - y\|^2 \right] + \lambda_1 \mathcal{L}_{\text{proj}} + \lambda_2 \mathcal{L}_{\text{recons}}, \quad (17)$$

where λ_1 and λ_2 are positive hyperparameters. In the first term, the squared error of the data point y^* and the prediction result y can be computed by solving the neural ODE represented as the equation (2).

The second term $\mathcal{L}_{\text{proj}}$ prevents the distance before and after projection from diverging by reducing a degree of freedom of projection, that is,

$$\begin{aligned} \mathcal{L}_{\text{proj}} &\triangleq \mathbb{E} \left[\|(\text{id}_{\mathbb{S}} - \mathcal{P}_{\theta})(f, g, h, j)\|^2 \right] \\ &= \mathbb{E}_{x \sim \mathcal{N}} \left[\|f(x) - f_{\mathbf{d}}(x)\|^2 + \|g(x) - g_{\mathbf{d}}(x)\|^2 \right], \end{aligned}$$

where $\text{id}_{\mathbb{S}}$ is the identity map on the set of dynamics \mathbb{S} . In our implementation of Theorem 6, only f and g are involved in the projection. Hence, $\mathcal{L}_{\text{proj}}$ can be calculated by the last formula, where the expectation by sampling from an n -dimensional standard normal distribution \mathcal{N} . In the following experiments, we use 100 samples to compute this expected values. Here, we emphasize that this loss term is used merely to reduce the degrees of freedom, even if its value is non-zero, our projected dynamics are always dissipative.

The last term is to prevent h from becoming degenerate in the early stages of gradient-based learning, that is,

$$\mathcal{L}_{\text{recons}} \triangleq \mathbb{E}_{x \in \mathcal{X}} \left[\|x - \eta(h(x))\| \right],$$

where $\eta : \mathbb{R}^l \rightarrow \mathbb{R}^n$ is represented by an additional neural network and \mathcal{X} is a set of $\eta(x)$ corresponding to y in the first term. where the reconstruction map $\eta : \mathbb{R}^l \rightarrow \mathbb{R}^n$ is represented by an additional neural network and \mathcal{X} is a set of samples on the solution of the neural ODE when computing the first term. The reconstruction term is intended to prevent h from becoming a trivial function $h(x) \equiv 0$ during learning. In our gradient-based learning, we initialize the neural network weights to values close to 0, and start learning from internal-state behaviors $x(t)$ around 0. In that case, h is learned as $h(x) \equiv 0$ at an early stage and does not change. The reconstruction term is intended to prevent this, empirically.

In this study, the map ℓ in the dissipative projection, the reconstruction map η , and nonlinear dynamics (f, g, h) are

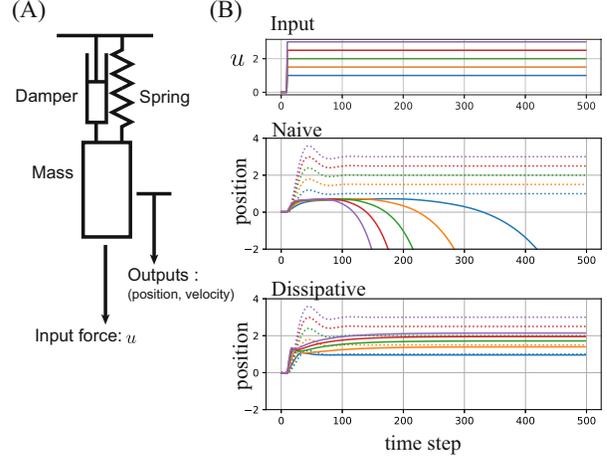


Figure 3: (A) Sketch of mass-spring-damper system. (B) Prediction results for out-of-domain inputs, i.e., five long step signals with different amplitudes. The top figure is the input signal behaviors, the middle figure shows the position of the mass predicted by the model trained using the naive model, and the bottom figure shows the position predicted by the proposed dissipative model. The dashed lines are the plots of the ground truth, and each color of the lines shows the results with the same input signals.

parameterized by using neural networks. All of the neural networks are trained using this loss function (17).

5 Result

We conducted three experiments to evaluate our proposed method. The first experiment uses a benchmark dataset generated from a mass-spring-damper system, which is a classic example from physics and engineering. In the next experiment, we evaluate our methods by an n -link pendulum system, a nonlinear dynamical system related to robotic arm applications. Finally, we applied our method to learning an input-output fluid system using a fluid simulator.

We evaluate the prediction error at each time point using root mean square error, which we call $\text{RMSE}(t)$, in the output domain. The aim is to observe the error affected by satisfying the dissipative system over time. In addition, we used the time-averaged $\text{RMSE}(t)$, which we refer to simply as RMSE , to evaluate the prediction error of models. In the following experiments, we performed experiments on the trained model by changing the input, such as test input sig-

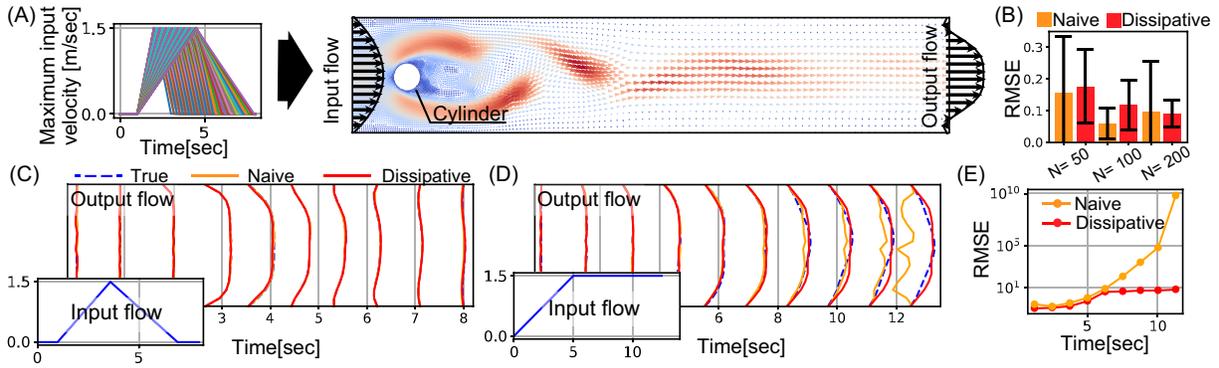


Figure 5: (A) Sketch of the flow around cylinder model. (B) Time-average RMSE of test triangular (in-domain) waves by changing the number of training inputs N . (C) The predicted output flow for a test triangular (in-domain) wave with $N = 100$. Each curved line represents the spatial distribution of the output flow at each time point. The blue, orange, and red lines show the ground truth, the output flow predicted by the naive model, and the output predicted by our dissipative model, respectively. (D) The predicted output flow for a test clip (out-of-domain) wave with $N = 100$. (E) $\text{RMSE}(t)$ for long time simulation with test clip waves.

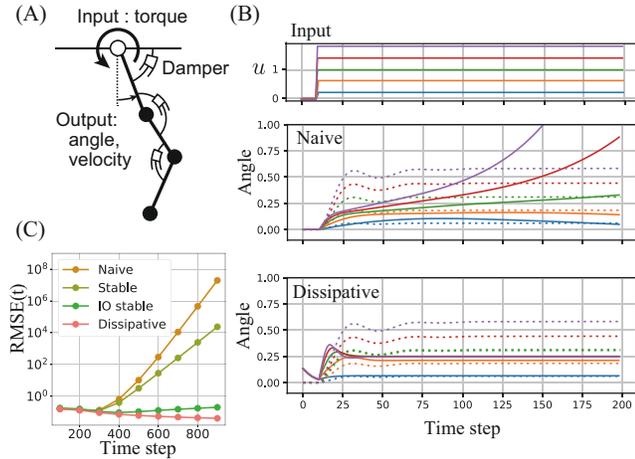


Figure 4: (A) Sketch of n -link pendulum model. (B) Prediction results for out-of-domain inputs. The top figure is the input signal behaviors, the middle figure shows the angle of the first pendulum predicted by the model trained using the naive model, and the bottom figure shows the angle predicted by the proposed dissipative model. The dashed lines are the plots of the ground truth, and each color of the lines shows the results with the same input signals. (C) $\text{RMSE}(t)$ related to long input step signal.

nal separated from training data, signals with different input lengths, and signals of different types, using these evaluation metrics. In the following experiment section, N denotes the number of pairs of input-output signals for training.

We retry five times for all experiments and show the mean and standard deviations of the both metrics. For simplicity in our experiments, the sampling step Δt for the output y is set as constant and the Euler method is used to solve neural ODEs. The initial state x_0 in this ODE is fixed as 0 for simplicity. The hyperparameters including the number of lay-

ers in the neural networks, the learning rate, optimizer, and the weighted decay are determined using the tree-structured Parzen estimator (TPE) implemented in Optuna (Akiba et al. 2019) (see Appendix L).

5.1 Mass-Spring-Damper Benchmark

We generate input and output signals for this experiment by a mass-spring-damper system. This dynamics was chosen as the first simple example because it is linear and its properties can be easily understood analytically.

Table 1 shows predictive performance comparing a naive model, the proposed stable model (Corollary 7), input-output stable model (Corollary 8), energy conservation model (Theorem 9), and dissipative model (Theorem 6). The naive model simply use neural networks as (f, g, h) , which is trained by minimizing the squared error. The first and fourth rows of this table shows the results of evaluation using N -inputs rectangle signals for training data and different $0.1 \times N$ rectangle input signals for testing.

Since our focus is on input-output systems, the model may be influenced by the type of input. Therefore, we considered a scenario where only data from input rectangle waves could be collected during training and evaluated the model by input signals generated by step functions and random walks, in addition to rectangle waves, during testing. We call such scenarios out-of-domain cases, shown in Table 1. The hyperparameters related to these models and the detail of the experimental setting are described in the Appendix K.1.

The proposed conservation and dissipative models exhibited high predictive performance with unforeseen inputs at $N = 100$. In particular, conservation showed a statistically significant improvement over the naive model. This is because the conservation model utilizes the energy relationships in the most rigorous manner. Note that the supply rate w includes not only the terms corresponding to increases in internal energy due to external forces but also the terms corresponding to energy dissipation by dampers. When the data size was increased to 1000, predictive accuracy differ-

ences between the methods were small. These observations suggest that enforcing dissipative or conservation properties ensures high predictive performance, particularly with smaller data sizes, due to better match with the underlying data-generating system.

Figure 3 illustrates another out-of-domain case when unexpectedly longer step inputs (1000 steps), exceeding those used during training (100 steps), are given. The results show that while the naive method may diverge with such an unexpected long input, ensuring dissipative properties allows the output to remain bounded.

5.2 n -Link Pendulum Benchmark

Next, to demonstrate the nonlinear case, we adopt the n -link pendulum system, characterized by multiple connected pendulums. The movement of each link is governed by nonlinear equations of motion, leading to extremely complex overall system behavior (Figure 4 (A)).

Figure 4 (B) and (C) show the behavior of the system when an input different from the 100-step rectangle waves used during training is input. The results show that divergence is evident from around 200 steps, indicating that naive and stable models may diverge when receiving an unexpectedly long input, but that the output is appropriately bounded when IO stability and dissipative are guaranteed. All numerical experiment results, including errors for the same type of input as used during training, are listed in the Appendix K.2.

5.3 Fluid System Benchmark

In the final part of this study, we aim to predict the input-output relationship of fluid flow around a cylinder (Figure 5 (A))(Schäfer et al. 1996). This phenomenon involves complex behaviors, such as periodic oscillations and flow instabilities, resulting from the formation of Kármán vortex streets. In this experiment, the left and right flow velocities are spatially discretized into 16 divisions, which are the inputs and outputs of this system.

We constructed a predictive model that guarantees dissipativity based on the results of fluid simulations using a triangular wave input (detailed conditions are provided in the Appendix K.3). The prediction results with test triangular wave inputs showed good accuracy for $N = \{50, 100, 200\}$, and a naive neural network also demonstrated comparable accuracy (Figure. 5 (B),(C)). Using out-of-domain clipped wave inputs, we compared the trained predictive model against long-term simulation. The results showed that the model maintained good predictive accuracy even for extended prediction periods, outperforming the naive model. The $RMSE(t)$ begins to increase at the time matching the training signal length (8 seconds) and becomes more pronounced over longer input signal (Figure. 5 (D), (E)).

6 Conclusion

In this study we analytically derived a general solution to the nonlinear KYP lemma and a dissipative projection. Furthermore, we showed that our proposed methods that strictly guarantee dissipativity including internal stability, input-output stability, and energy conservation. Finally, we con-

firmed the effectiveness of our method, particularly its robustness to out-of-domain input, using both linear systems and complex nonlinear systems, including an n -link pendulum and fluid simulations. A limitation of this study is the requirement to determine the dissipative hyperparameters based on the rough properties of the target system. For future work, this research will lead to system identification to control real-world dissipative systems.

Acknowledgments

This research was supported by JST Moonshot R&D Grant Number JPMJMS2021 and JPMJMS2024. This work was also supported by JSPS KAKENHI Grant No.21H04905 and CREST Grant Number JPMJCR22D3, Japan.

References

- Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; and Koyama, M. 2019. Optuna: A Next-generation Hyperparameter Optimization Framework. In *Proceedings of the 25rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.
- Blocher, C.; Saveriano, M.; and Lee, D. 2017. Learning stable dynamical systems using contraction theory. In *2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, 124–129. IEEE.
- Brogliato, B.; Lozano, R.; Maschke, B.; and Egeland, O. 2020. *Dissipative systems analysis and control: Theory and Applications*. Springer, 3rd edition.
- Chang, Y.-C.; Roohi, N.; and Gao, S. 2019. Neural Lyapunov Control. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 32.
- Chen, R. T.; Rubanova, Y.; Bettencourt, J.; and Duvenaud, D. 2018. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 31.
- Chen, X. 2019. Review: Ordinary Differential Equations For Deep Learning. *CoRR*, abs/1911.00502.
- Drgoňa, J.; Tuor, A.; Vasiht, S.; and Vrabie, D. 2022. Dissipative deep neural dynamical systems. *IEEE Open Journal of Control Systems*, 1: 100–112.
- Goldbeter, A. 2018. Dissipative structures in biological systems: bistability, oscillations, spatial patterns and waves. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2124): 20170376.
- Greydanus, S.; Dzamba, M.; and Yosinski, J. 2019. Hamiltonian neural networks. *Advances in neural information processing systems*, 32.
- Hatanaka, T.; Chopra, N.; Fujita, M.; and Spong, M. W. 2015. *Passivity-based control and estimation in networked robotics*. Springer.
- Khansari-Zadeh, S. M.; and Billard, A. 2011. Learning stable nonlinear dynamical systems with gaussian mixture models. *IEEE Transactions on Robotics*, 27(5): 943–957.

Kidger, P.; Morrill, J.; Foster, J.; and Lyons, T. 2020. Neural controlled differential equations for irregular time series. *Advances in Neural Information Processing Systems*, 33: 6696–6707.

Kojima, R.; and Okamoto, Y. 2022. Learning deep input-output stable dynamics. *Advances in Neural Information Processing Systems*, 35: 8187–8198.

Lawrence, N.; Loewen, P.; Forbes, M.; Backstrom, J.; and Gopaluni, B. 2020. Almost Surely Stable Deep Dynamics. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 33.

Manek, G.; and Zico Kolter, J. 2019. Learning stable deep dynamics models. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 32.

Morrill, J.; Salvi, C.; Kidger, P.; and Foster, J. 2021. Neural rough differential equations for long time series. In *International Conference on Machine Learning*, 7829–7838. PMLR.

Ortega, R.; and Ortega, R. 1998. *Passivity-based control of Euler-Lagrange systems : mechanical, electrical, and electromechanical applications*. Communications and Control Engineering. London, England: Springer-Verlag, 1st ed. 1998. edition. ISBN 1-4471-3603-9.

Richards, S. M.; Berkenkamp, F.; and Krause, A. 2018. The lyapunov neural network: Adaptive stability certification for safe learning of dynamical systems. In *Conference on Robot Learning*, 466–476. PMLR.

Schäfer, M.; Turek, S.; Durst, F.; Krause, E.; and Rannacher, R. 1996. *Benchmark Computations of Laminar Flow Around a Cylinder*, 547–566. Wiesbaden: Vieweg+Teubner Verlag. ISBN 978-3-322-89849-4.

Sosanya, A.; and Greydanus, S. 2022. Dissipative hamiltonian neural networks: Learning dissipative and conservative dynamics separately. *arXiv preprint arXiv:2201.10085*.

Stramigioli, S. 2006. Geometric control of mechanical systems: modelling, analysis, and design for simple mechanical control systems, Francesco Bullo and Andrew D. Lewis, Springer, New York, NY, 2005. *International journal of robust and nonlinear control*, 16(11): 547–548. 10.1002/rnc.1064.

Takeishi, N.; and Kawahara, Y. 2021. Learning Dynamics Models with Stable Invariant Sets. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, 9782–9790.

Umlauft, J.; and Hirche, S. 2017. Learning Stable Stochastic Nonlinear Dynamical Systems. In Precup, D.; and Teh, Y. W., eds., *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, 3502–3510. PMLR.

Xu, Y.; and Sivaranjani, S. 2023. Learning Dissipative Neural Dynamical Systems. *IEEE Control Systems Letters*, 7: 3531–3536.