ORIGINAL PAPER

Port-metriplectic neural networks: thermodynamics-informed machine learning of complex physical systems

Quercus Hernández[1] · Alberto Badías[2] · Francisco Chinesta[3,4] · Elías Cueto[1]

Abstract
We develop inductive biases for the machine learning of complex physical systems based on the port-Hamiltonian formalism. To satisfy by construction the principles of thermodynamics in the learned physics (conservation of energy, non-negative entropy production), we modify accordingly the port-Hamiltonian formalism so as to achieve a port-metriplectic one. We show that the constructed networks are able to learn the physics of complex systems by parts, thus alleviating the burden associated to the experimental characterization and posterior learning process of this kind of systems. Predictions can be done, however, at the scale of the complete system. Examples are shown on the performance of the proposed technique.

Keywords Port-Hamiltonian · Thermodynamics · Scientific machine learning · Inductive biases

1 Introduction

Recently, the possibility of developing learned simulators has attracted an important research activity in the computational mechanics community and beyond. By "learned simulators" we mean methodologies able to learn from data the dynamics of a physical system so as to perform accurate predictions about previously unseen situations without the burden asso-

Quercus Hernández and Alberto Badías have contributed equally to this work.

✉ Elías Cueto
  ecueto@unizar.es

  Quercus Hernández
  quercus@unizar.es

  Alberto Badías
  alberto.badias@upm.es

  Francisco Chinesta
  Francisco.Chinesta@ensam.eu

[1] Aragon Institute of Engineering Research, Universidad de Zaragoza, Maria de Luna, s.n., 50018 Zaragoza, Spain

[2] Higher Technical School of Industrial Engineering, Polytechnic University of Madrid, C. de José Gutiérrez Abascal, 2, 28006 Madrid, Spain

[3] ESI Group Chair, PIMM Lab, ENSAM Institute of Technology, 151 Bvd. de l'Hôpital, 75013 Paris, France

[4] CNRS@CREATE LTD., CNRS, 1 CREATE Way, Singapore 138602, Singapore

ciated to the construction of numerical models by means of finite elements, finite volumes or similar techniques [1–3]. Among their advantages we can cite that they are based on reusable architectures, can be optimized to work under really stringent real-time feedback rates, and are specially well suited for optimization and inverse problems.

While original, black-box approaches showed great promise, both industry and academia are reluctant to generalize their use, since small modifications in the input data may cause nonsense results. This is at the origin of the development and employ of inductive biases during the learning process [3,4]. An inductive bias allows the learning algorithm to prioritize one particular solution over any other [5]. This is particularly interesting for physical phenomena for which previous knowledge exists. Paul Dirac once said that [6]

> *The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.*

Therefore, in the presence of centuries of knowledge about virtually any physical phenomena, it is simply nonsense to ignore it and to favor theory-blind, black-box approaches.

In this paper we develop a novel strategy based on the port-Hamiltonian formalism, which we extend so as to comply with the first and second principles of thermodynamics by construction [7–9]. Port-Hamiltonian formalisms extend

the well-known Hamiltonian (thus, conservative) physics to open systems and introduce the possibility of dissipation and control through external actuation within this theory. We show here, however, that general port-Hamiltonian systems do not comply a priori with the laws of thermodynamics and modify them so as to ensure this fulfillment. Based on this new formalism, which we call port-metriplectic, since it is at the same time metric and symplectic, we construct a deep neural network methodology to learn the physics of complex systems from data. The resulting port-metriplectic networks will comply by construction with the principles of thermodynamics—that can be enforced through hard or soft constraints—while they allow to analyze complex systems by parts. These parts will then communicate through energy ports to construct the final, complex systems.

The outline of the paper is as follows. In Sect. 2 we review the state of the art in the development of machine learning strategies that impose energy conservation by using a Hamiltonian formalism. We include here neural networks based upon port-Hamiltonian formalisms, which we show not be necessarily compliant with the principles of thermodynamics. We then develop the concept of port-metriplectic networks in Sect. 3. Then, in Sect. 4 we analyze the performance of the just developed neural networks, while in Sect. 5 we draw some conclusions.

2 Hamiltonian neural networks

2.1 Reversible dynamics as an inductive bias

Learning the physics of a given phenomenon from data can be seen as learning a dynamical problem [10]. If we assume that the problem is governed by a set of variables z, which we can measure experimentally—a detailed discussion on the limitations and implications of this assumption can be found at [11], then the problem of learning the evolution of the system in time can be seen as finding the structure of the dynamical problem

$$\dot{z} = \frac{dz}{dt} = f(z, t), \quad z(0) = z_0, \tag{1}$$

or, in other words, to find by regression the flow map

$$z_0 \rightarrow z(T, z_0). \tag{2}$$

Equivalently, we must find the particular form of the function f governing the dynamics of the system. This is done by regression, where neural networks play an important role, but by no means constitute the only possibility, as in [12–14].

This particular form of seeing the problem has important advantages. For instance, if the system under scrutiny is known to be conservative, or reversible, we can impose as an inductive bias the Hamiltonian form of the sought function f,

$$\dot{z} = \frac{dz}{dt} = L\frac{\partial \mathcal{H}}{\partial z} = L\frac{\partial E}{\partial z}, \tag{3}$$

where the Hamiltonian, $\mathcal{H}$, whose canonical form depends on the position and momenta of particles, is now the total energy of the system, E. Under this prism, the problem (1) is now seen as to find the precise form of the skew-symmetric (symplectic) matrix L and the form of the energy of the system, $E(z)$. If we enforce the particular form given by Eq. (3) during our regression procedure, it is straightforward to prove that the resulting evolution will be conservative.

Many works have leveraged this approach. Several authors take advantage of the Hamiltonian structure to construct symplectic integrators to predict conservative dynamical systems [15–17]. Others, use the Hamiltonian principles to design more expressive deep neural network architectures [18] or to find the Hamiltonian function and phase space from data [19,20]. The Hamiltonian paradigm is also widely used in quantum mechanics, where similar deep learning literature can be found in problems such as electron dynamics [21], learning ground states [22] or optimal control [23]. Alternative formulations can be developed by resorting to the equivalent Lagrangian formalism, see [4,24–27], among others.

2.2 Port-Hamiltonian neural networks

If the physical phenomenon at hand is known to be dissipative, or if the system is open and thus no guarantee on the conservation of energy exists, things become more intricate. For dissipative systems, the easiest form of the evolution Eq. (1) could be, perhaps, a gradient flow [28]. Their evolution can be established after some (dissipation) potential $\mathcal{R}$ in the form [29]

$$\dot{z} = -\frac{\partial \mathcal{R}}{\partial z}.$$

Recently, the so-called Symplectic ODE nets (symODEN) [30,31], have tackled the issue of introducing dissipation in the learned description of the physics. It is also the approach followed in [32]. More recently, two distinct works have tackled the dissipation problem by relaxing equivariance in the networks [33,34].

These works seem to be closely related to the vast corps of literature on the port-Hamiltonian approach to dynamical systems [7–9]. Port-Hamiltonian systems assume an evolution of the system in the form

$$\begin{bmatrix} \dot{q} \\ \dot{p} \end{bmatrix} = \left(\begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{bmatrix} - D(q) \right) \begin{bmatrix} \frac{\partial \mathcal{H}}{\partial q} \\ \frac{\partial \mathcal{H}}{\partial p} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ g(q) \end{bmatrix} u, \tag{4}$$

where (q, p) are the generalized position and momenta, dissipation is included by adding a symmetric, positive semi-definite matrix D, and control is considered through an actuation term u and a non-linear function of the position $g(q)$. Equation (4) reduces to the Hamiltonian description if no dissipation nor control are considered. Here, we have assumed a canonical form for the Hamiltonian, i.e., that it depends on a set of variables $z = \{q, p\}$. More general forms can be expressed similarly.

The true advantage of using port-Hamiltonian formalisms as inductive biases in the learning procedure stems from the fact that, on one side, they allow the introduction of dissipation and control and, on the other, they model open systems (as opposed to classical Hamiltonian descriptions where energy conservation assumes inherently that the system is closed) [35].

Therefore, the use of port-Hamiltonian formalisms as inductive biases in learning processes is extremely interesting. However, as will be demonstrated in the next section, classical port-Hamiltonian schemes do not guarantee a priori to comply with the laws of thermodynamics, see ref. [11].

3 Port-metriplectic neural networks

3.1 Metriplectic biases for dissipative phenomena

In the case of dissipative phenomena, the first in proposing the introduction of a second potential, the so-called Mathieu potential, seems to have been Morrison [36,37], Grmela [38,39] and Kaufman [40]. They suggested to consider an evolution of the governing variables of the type

$$\dot{z} = L(z) \frac{\partial E}{\partial z} + M(z) \frac{\partial S}{\partial z}, \tag{5}$$

where S is precisely this second (dissipation) potential, entropy.

This formulation is often referred to as metriplectic, since it is metric and symplectic at the same time. Here, $M(z)$ is a symmetric, positive semi-definite dissipation matrix and $L(z)$, the Poisson matrix, continues to be skew-symmetric.

However, for this formulation to be consistent with the principles of thermodynamics, two additional conditions must hold, the so-called degeneracy conditions:

$$L(z) \frac{\partial S}{\partial z} = \mathbf{0}, \tag{6}$$

and

$$M(z) \frac{\partial E}{\partial z} = \mathbf{0}, \tag{7}$$

which give rise to the General Equation for the non-Equilibrium Reversible-Irreversible Coupling, GENERIC, equations [41–45].

In a nutshell, Eqs. (6) and (7) state that the energy potential is independent of dissipation, whereas entropy is unrelated to the energy conservation. If they hold, it is straightforward to demonstrate that, given the skew-symmetry of L,

$$\dot{E}(z) = \frac{\partial E}{\partial z} \dot{z} = 0,$$

and

$$\dot{S} = \frac{\partial S}{\partial z} \dot{z} = \frac{\partial S}{\partial z} M(z) \frac{\partial S}{\partial z} \geq 0,$$

given the positive semi-definiteness of M.

These properties have been leveraged in some of our former works to develop what we have coined as thermodynamics-informed neural networks [46–48].

Given experimental data sets $\mathcal{D}_i$ containing labelled pairs of a single-step state vector z_t and its evolution in time z_{t+1},

$$\mathcal{D} = \{\mathcal{D}_i\}_{i=1}^{N_{\text{sim}}}, \quad \mathcal{D}_i = \{(z_t, z_{t+1})\}_{t=0}^{T}, \tag{8}$$

we construct a neural network by considering two different loss terms. First, a data-loss term that takes into account the correctness of the network prediction of the state vector at subsequent time steps by integrating GENERIC in time, i.e.,

$$\mathcal{L}_n^{\text{data}} = \left\| \dot{z}_n^{\text{GT}} - \dot{z}_n^{\text{net}} \right\|_2^2,$$

with $\| \cdot \|_2$ the L2-norm, $\dot{z}_n^{\text{GT}}$ is ground truth solution and $\dot{z}_n^{\text{net}}$ is the network prediction. The choice of the time derivative instead of the state vector itself is employed to regularize the global loss function to a uniform order of magnitude with respect to the degeneracy terms.

We then consider a second loss term to take into account the fulfillment of the degeneracy equations in a soft way,

$$\mathcal{L}_n^{\text{deg}} = \left\| L \frac{\partial S}{\partial z_n} \right\|_2^2 + \left\| M \frac{\partial E}{\partial z_n} \right\|_2^2.$$

These networks have have demonstrated to work very well for physics perception and reasoning in combination with computer vision [49,50].

These two loss terms are weighted and averaged over the N_{batch} batched snapshots.

$$\mathcal{L} = \frac{1}{N_{\text{batch}}} \sum_{n=0}^{N_{\text{batch}}} (\lambda \mathcal{L}_n^{\text{data}} + \mathcal{L}_n^{\text{deg}}). \tag{9}$$

Alternative formulations of these thermodynamics-informed networks exist in which the degeneracy conditions are imposed in hard way, see refs. [51] and [52].

It is worth noting that, by comparing Eqs. (5), (6) and (7), on one side, and Eq. (4), on the other, one readily concludes that port-Hamiltonian biases do not necessarily ensure the fulfillment of the principles of thermodynamics. Note that, since entropy does enter the classical port-Hamiltonian formulation, it is difficult to impose the fulfillment of the second principle of thermodynamics. Therefore, we suggest to extend the GENERIC formalism to open systems so as to develop alternative port-metriplectic biases. These are developed in the next section.

3.2 Port-metriplectic neural networks

Very few works exist, to the best of our knowledge, on the development of GENERIC formulations for open systems, that may lead to the development of port-metriplectic formulations. Maybe the only exception is [53], later on revisited by [54,55], both published in conference proceedings and, of course, with no machine learning approximations. Both approaches are essentially identical, and start from the bracket formulation of GENERIC. Operators L and M define a bracket structure of the type

$$\dot{z} = \{z, E\} + [z, S], \tag{10}$$

where $\{\cdot, \cdot\}$ is the so-called Poisson bracket and $[\cdot, \cdot]$ represents the dissipative bracket [41,45,56].

For open systems, these brackets take the form

$$\{\cdot, \cdot\} = \{\cdot, \cdot\}_{\text{bulk}} + \{\cdot, \cdot\}_{\text{boun}}, \tag{11}$$

and

$$[\cdot, \cdot] = [\cdot, \cdot]_{\text{bulk}} + [\cdot, \cdot]_{\text{boun}}. \tag{12}$$

In other words, both brackets are decomposed additively into bulk and boundary contributions. With this decomposition in mind, the GENERIC principle (5) now reads

$$\begin{aligned} \dot{z} &= \{z, E\}_{\text{bulk}} + [z, S]_{\text{bulk}} \\ &= \{z, E\} + [z, S] - \{z, E\}_{\text{boun}} - [z, S]_{\text{boun}}. \end{aligned} \tag{13}$$

The degeneracy conditions (6) and (7) must be satisfied by the bulk operators only, since it is possible, in general, that

there may be a reversible flux of entropy at the boundary or, equivalently, an irreversible flux of energy at the boundary [53],

$$L_{\text{bulk}}(z) \frac{\partial S_{\text{bulk}}}{\partial z} = \mathbf{0}, \tag{14}$$

and

$$M_{\text{bulk}}(z) \frac{\partial E_{\text{bulk}}}{\partial z} = \mathbf{0}, \tag{15}$$

The particular form of the boundary terms in Eq. (13) depends, of course, of the particular phenomenon under scrutiny, but in a general way it can be expressed using L and M operators as

$$\begin{aligned} \dot{z} &= L \frac{\partial E}{\partial z} + M \frac{\partial S}{\partial z} \\ &\quad - L_{\text{boun}} \frac{\partial E_{\text{boun}}}{\partial z} - M_{\text{boun}} \frac{\partial S_{\text{boun}}}{\partial z}. \end{aligned} \tag{16}$$

More particular expressions can be developed if we know in advance some properties of the system at hand. For instance, in Sect. 4.1 we deal with a double pendulum by learning the behavior of each pendulum separately. If we know in advance that the only boundary term comes from the energy-entropy pair transmitted by the other pendulum, and no other external contribution is present, more detailed assumptions in the form of degeneracy conditions can be assumed. This may lead to a decrease in learning time or the employ of less data.

Figure 1 sketches the approach developed herein for complex systems. In the numerical results section below we explore the particular form that these terms could acquire for both finite and infinite dimensional problems.

We propose two learning procedures which correspond to different level of information available of the dynamics of the system. In the first example, we focus on two coupled subsystems in which we learn the self and boundary contributions of both subsystems to the global dynamics of the problem. This is the case when the interest is focused on the complete system divided into smaller subsystems. In the second example, we suppose that the external influence is determined by a load vector as a result of an unknown external interaction with another subsystem. Thus, the learning procedure is focused on the self and boundary contributions of only one subsystem based on an external interaction. This case is convenient for applications where only partial information of the system is available.
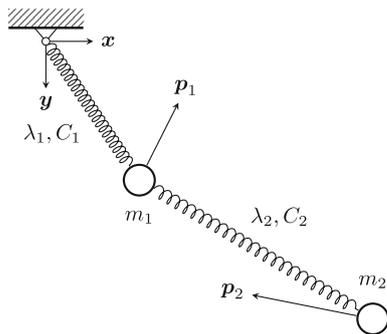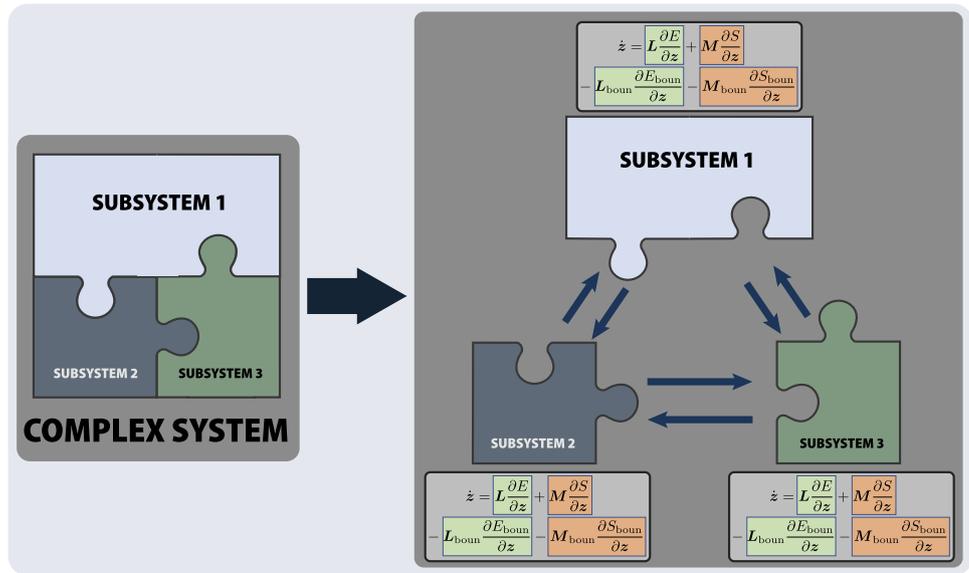
Fig. 2 Double thermoelastic pendulum. A single pendulum with external perturbations is learned, and the coupling between both systems is achieved via the port-metriplectic framework

4 Numerical results

4.1 Double thermoelastic pendulum

The first example is a double thermoelastic pendulum consisting of two masses m_1 and m_2 connected by two springs of variable lengths λ_1 and λ_2 and natural lengths at rest λ_1^0 and λ_2^0, as depicted in Fig. 2.

The set of variables describing each pendulum are here chosen to be

$$\mathcal{S} = \{z = (q, p, s) \in (\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R})\}. \tag{17}$$

where q, p and s are the position, linear momentum and entropy of the pendulum mass.

The evolution of the state variables of the second pendulum is defined as

$$\dot{z}_2 = L_2 \frac{\partial E_2}{\partial z_2} + M_2 \frac{\partial S_2}{\partial z_2} - M_{\text{boun},2} \frac{\partial S_{\text{boun},2}}{\partial z_2},$$

where the first two positive terms describe the self contribution of the simple pendulum (conservative and dissipative effects) and the third term describes the dissipative effect produced by the first pendulum affecting over the second pendulum.

On the other hand, the evolution of the state variables of the first pendulum is defined as

$$\dot{z}_1 = L_1 \frac{\partial E_1}{\partial z_1} + M_1 \frac{\partial S_1}{\partial z_1}$$
$$- L_{\text{boun},1} \frac{\partial E_{\text{boun},1}}{\partial z_1} - M_{\text{boun},1} \frac{\partial S_{\text{boun},1}}{\partial z_1}, \tag{18}$$

where in this case the first two positive terms describe the self contribution of the first simple pendulum (conservative and dissipative effects) and the third and fourth terms describe the external contribution on the conservative and dissipative parts, both produced by the influence of the second pendulum over the first pendulum.

Note that the first pendulum has no conservative contribution to the second pendulum, i.e., the term

$$L_{\text{boun},2} \frac{\partial E_{\text{boun},2}}{\partial z_2}$$

does not exist. However, there is a conservative contribution from the second pendulum on the first pendulum, see [57].

It is worth noting, as previously pointed out in [35], that the fact that every term in Eq. (18) depends on the state variables z_1 makes the learning procedure more intricate. This is caused by the non-separable structure of Eq. (18). This problem is not present if the port terms depend only on time, as it is the case in Sect. 4.2 below. To overcome this limitation, we employ a structure-preserving neural network for each of the terms in Eq. (18). These networks share the weights,
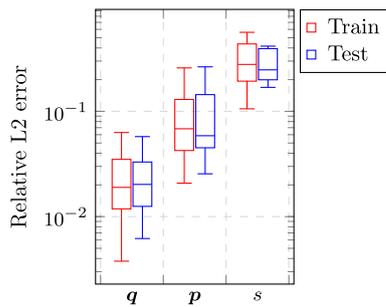
Fig. 3 Box plots for the relative L2 error for all the rollout snapshots of the double pendulum in both train and test cases. The state variables represented are position (q), momentum (p), and entropy (s)

however, for both pendula, if they are known in advance to be identical.

The fact of using individual approximations of the dynamics of each subsystem (each pendulum) allows to use artificial neural networks of considerably smaller size with respect to an analysis of the whole problem using a larger number of variables to describe the global state [46].

The database consists of 50 different simulations with random initial conditions of position q and linear momentum p of both masses m_1 and m_2 around a mean position and linear momentum of $q_1 = (4.5, \ 4.5)$ m, $p_1 = (2, \ 4.5)$ kg·m/s, and $q_2 = (-0.5, \ 1.5)$ m, $p_2 = (1.4, \ -0.2)$ kg·m/s respectively. The masses of the double pendulum are set to $m_1 = 1$ kg and $m_2 = 2$ kg, joint with springs of a natural length of $\lambda_1^0 = 2$ m and $\lambda_2^0 = 1$ m and thermal constant of $C_1 = 0.02$ J and $C_2 = 0.2$ J and conductivity constant of $\kappa = 0.5$. Note that the double pendulum constitutes a closed system as a whole, but this is not the case for each one of the simple pendula. Both start from a temperature of 300 K. The simulation time of the movement is $T = 60$ s in $N_T = 200$ time increments of $\Delta t = 0.3$ s.

The boxplot in Fig. 3 shows the statistics of the L2 relative error of the rollout train and test simulations.

4.2 Interacting beams

In this example we consider two viscoeleastic beams that can interact through contact between them, see Fig. 4, and whose physics are to be learned. Synthetic data come from finite element simulations, assuming a strain energy potential of the type

$$U = C_{10}(\bar{I}_1 - 3) + C_{01}(\bar{I}_2 - 3) + \frac{1}{D_1}(J_{el} - 1)^2,$$

with J_{el} the elastic volume ratio, $\bar{I}_1$ and $\bar{I}_2$ are the two invariants of the left Cauchy-Green deformation tensor, C_{10} and C_{01} are shear material constants and D_1 is the material compressibility parameter. The viscoelastic behavior is described
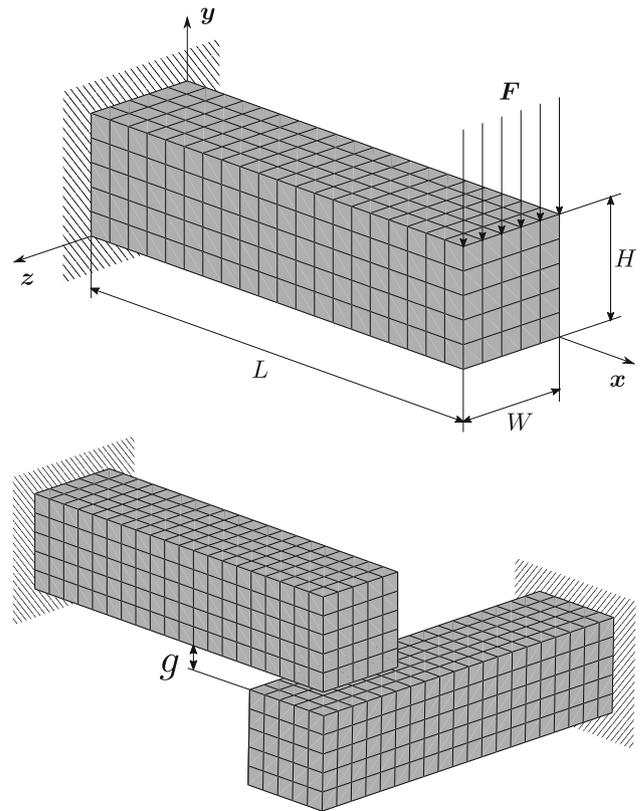
Fig. 4 One single beam problem is analyzed from simulation data. The resulting system is composed of two of these beams interacting together

by a two-term Prony series of the dimensionless shear relaxation modulus,

$$g_R(t) = 1 - \bar{g}_1(1 - e^{\frac{-t}{\tau_1}}) - \bar{g}_2(1 - e^{\frac{-t}{\tau_2}}),$$

with relaxation coefficients of $\bar{g}_1$ and $\bar{g}_2$, and relaxation times of τ_1 and τ_2.

We assume that the necessary state variables for for a proper description of the beams are the position q, its velocity v and the stress tensor σ,

$$S = \{z = (q, v, \sigma) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^6\}, \qquad (19)$$

at each node of the discretization of the beams. Since both beams are identical, see Fig. 4 we characterize only one of them and develop a port-metriplectic learned simulator for the joint system. To do so, we employ thermodynamics-informed graph neural networks [48].

Basically, a graph neural network is constructed on top of a graph structure $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \boldsymbol{u})$, where $\mathcal{V} = \{1, \ldots, n\}$ is a set of $|\mathcal{V}| = n$ vertices, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is a set of $|\mathcal{E}| = e$ edges and $\boldsymbol{u}$ is the global feature vector. Each vertex and edge in the graph is associated with a node in the finite element model from which data are obtained. The global feature vec-
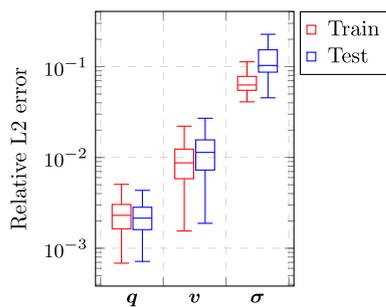
Fig. 5 Box plots for the relative L2 error for all the rollout snapshots of the interacting beams in both train and test cases. The state variables represented are position (q), velocity (v), and stress tensor (σ)

tor defines the properties shared by all the nodes in the graph, such as constitutive properties. More details on the precise formulation can be found at [48].

To ensure traslational invariance of the learned model, the position variables of the system, q_i, are assigned to the edge feature vector e_{ij} so the edge features represent relative distances ($q_{ij} = q_i - q_j$) between nodes. The rest of the state variables are assigned to the node feature vector v_i. We employ an encode-process-decode scheme [3], built upon multilayer perceptrons (MLPs) shared between all the nodes and edges of the graph.

We use this graph-based framework to learn the self contribution of the dynamics, i.e. the first two terms of Eq. (16). The boundary terms are learned using a standard structure-preserving neural network [46] with the additional input of the external forces applied to the beam.

The dimensions of the beams are $H = 10$, $W = 10$ and $L = 40$. The finite element mesh from which data are obtained consisted of $N_e = 500$ hexahedral linear brick elements and $N = 756$ nodes. The constitutive parameters are $C_{10} = 1.5 \cdot 10^5$, $C_{01} = 5 \cdot 10^3$, $D_1 = 10^{-7}$ and $\bar{g}_1 = 0.3$, $\bar{g}_2 = 0.49$, $\tau_1 = 0.2$, $\tau_2 = 0.5$ respectively. A distributed load of $F = 10^5$ is applied in 52 different positions with an orientation perpendicular to the solid surface. Simulations were quasi-static and included $N_T = 20$ time increments of $\Delta t = 5 \cdot 10^{-2}$ s. Two identical beams are assembled in 90° with a gap of $g = 10$, as depicted in Fig. 4.

The results are presented in Fig. 5. The error magnitude is similar as the reported in previous work [48] in addition to the consistent formulation of port-metriplectic dynamics.

classical port-Hamiltonian approaches while guaranteeing the fulfillment of the laws of thermodynamics (conservation of energy in the bulk system, non-negative entropy production). The resulting formulation, which we refer as port-metriplectic—since it consists of a metric term and a symplectic one—presents a rigorous thermodynamic description of the dissipative behavior of the system.

On the other hand, the just developed formulation is employed as an inductive bias for the machine learning of the physics of complex systems from measured data. This bias is developed as a soft constraint in the loss term, although it can also be imposed straightforwardly as a hard constraint.

The resulting neural networks, for which we have formulated two distinct versions, one based on standard multilayer perceptrons, and a second one based on graph neural networks, have shown an excellent performance. Error bars are equivalent to those obtained in previous works of the authors, by employing a closed-system approach to the same physics. The new approach opens the door to the development of learned simulators for complex systems through piece-wise learning of the physical behavior of each of its components. The final, global simulator is then obtained by assembling each piece through their ports.

5 Conclusions

In this paper we have made two main contributions. On one side, the development of port-Hamiltonian-like approximations for dissipative open systems that communicate with other systems by exchanging energy and entropy through ports in their boundaries. This formulation extends the

References

1. Stachenfeld K, Fielding DB, Kochkov D, Cranmer M, Pfaff T, Godwin J, Cui C, Ho S, Battaglia P, Sanchez-Gonzalez A (2021) Learned simulators for turbulence. In: International conference on learning representations

2. Allen KR, Lopez-Guevara T, Stachenfeld K, Sanchez-Gonzalez A, Battaglia P, Hamrick J, Pfaff T (2022) Physical design using differentiable learned simulators. arXiv preprint arXiv:2202.00728

3. Battaglia PW, Hamrick JB, Bapst V, Sanchez-Gonzalez A, Zambaldi V, Malinowski M, Tacchetti A, Raposo D, Santoro A, Faulkner R et al (2018) Relational inductive biases, deep learning, and graph networks. arXiv preprint arXiv:1806.01261

4. Bhattoo Ravinder, Ranu Sayan, Krishnan NM (2021) Lagrangian neural network with differentiable symmetries and relational inductive bias. arXiv preprint arXiv:2110.03266

5. Mitchell TM (1980) The need for biases in learning generalizations. Department of Computer Science, Laboratory for Computer Science Research

6. Dirac PAM (1929) Quantum mechanics of many-electron systems. Proc R Soc Lond Ser A Contain Papers Math Phys Character 123(792):714–733

7. Van Der Schaft A et al (2014) Port-Hamiltonian systems theory: an introductory overview. Found Trends® Syst Control 1(2–3):173–378

8. Beattie CA, Mehrmann V, Van DP (2019) Robust port-Hamiltonian representations of passive systems. Automatica 100:182–186

9. Rashad R, Califano F, van der Schaft AJ, Stramigioli S (2020) Twenty years of distributed port-hamiltonian systems: a literature review. IMA J Math Control Inf 37(4):1400–1422

10. Weinan E (2017) A proposal on machine learning via dynamical systems. Commun Math Stat 5(1):1–11

11. Cueto E, Chinesta F (2022) Thermodynamics of learning physical phenomena. arXiv preprint arXiv:2207.12749

12. González D, Chinesta F, Cueto E (2019) Thermodynamically consistent data-driven computational mechanics. Contin Mech Thermodyn 31(1):239–253

13. González D, Chinesta F, Cueto E (2019) Learning corrections for hyperelastic models from data. Front Mater 6:14

14. González D, Chinesta F, Cueto E (2021) Learning non-Markovian physics from data. J Comput Phys 428:109982

15. Jin P, Zhang Z, Zhu A, Tang Y, Karniadakis GE (2020) Sympnets: intrinsic structure-preserving symplectic networks for identifying Hamiltonian systems. Neural Netw 132:166–179

16. Chen Z, Feng M, Yan J, Zha H (2022) Learning neural Hamiltonian dynamics: a methodological overview. arXiv preprint arXiv:2203.00128

17. Miller ST, Lindner JF, Choudhary A, Sinha S, Ditto WL (2020) Mastering high-dimensional dynamics with Hamiltonian neural networks. arXiv preprint arXiv:2008.04214

18. Galimberti CL, Xu L, Trecate GF (2021) A unified framework for Hamiltonian deep neural networks. In: Learning for dynamics and control, pp 275–286. PMLR, USA

19. Bertalan T, Dietrich F, Mezić I, Kevrekidis IG (2019) On learning Hamiltonian systems from data. Chaos Interdiscip J Nonlinear Sci 29(12):121107

20. Toth P, Rezende DJ, Jaegle A, Racanière S, Botev A, Higgins I (2019) Hamiltonian generative networks. arXiv preprint arXiv:1909.13789

21. Bhat HS, Ranka K, Isborn CM (2020) Machine learning a molecular Hamiltonian for predicting electron dynamics. Int J Dyn Control 8(4):1089–1101

22. Kochkov D, Pfaff T, Sanchez-Gonzalez A, Battaglia P, Clark BK (2021) Learning ground states of quantum hamiltonians with graph networks. arXiv preprint arXiv:2110.06390

23. Yuanqi G, Wang X, Nanpeng Y, Wong BM (2022) Harnessing deep reinforcement learning to construct time-dependent optimal fields for quantum control dynamics. Phys Chem Chem Phys 24(39):24012–24020

24. Lutter M, Ritter C, Peters J (2019) Deep Lagrangian networks: using physics as model prior for deep learning. arXiv preprint arXiv:1907.04490

25. Zhong YD, Leonard N (2020) Unsupervised learning of Lagrangian dynamics from images for prediction and control. Adv Neural Inf Proc Syst 33:10741–10752

26. Lee S-C, Kim Y-H (2002) An enhanced Lagrangian neural network for the eld problems with piecewise quadratic cost functions and nonlinear constraints. Electr Power Syst Res 60(3):167–177

27. Allen-Blanchette C, Veer S, Majumdar A, Leonard NE (2020) Lagnetvip: a Lagrangian neural network for video prediction. arXiv preprint arXiv:2010.12932

28. Hohenberg PC, Halperin BI (1977) Theory of dynamic critical phenomena. Rev Modern Phys 49(3):435

29. Weinan E (2020) Machine learning and computational mathematics. Commun Comput Phys 28(5):1639–1670

30. Zhong YD, Dey B, Chakraborty A (2020) Dissipative symoden: encoding hamiltonian dynamics with dissipation and control into deep learning. arXiv preprint arXiv:2002.08860

31. Zhong YD, Dey B, Chakraborty A (2021) Benchmarking energy-conserving neural networks for learning dynamics from data. In: Learning for dynamics and control, pp 1218–1229. PMLR, USA

32. Gruver N, Finzi M, Stanton S, Wilson AG (2022) Deconstructing the inductive biases of hamiltonian neural networks. arXiv preprint arXiv:2202.04836

33. Han J, Huang W, Ma H, Li J, Tenenbaum JB, Gan C (2022) Learning physical dynamics with subequivariant graph neural networks. arXiv preprint arXiv:2210.06876

34. Wang R, Walters R, Yu R (2022) Approximately equivariant networks for imperfectly symmetric dynamics. arXiv preprint arXiv:2201.11969

35. Eidnes S, Stasik AJ, Sterud C, Bøhn E, Riemer-Sø RS (2022) Port-hamiltonian neural networks with state dependent ports. arXiv preprint arXiv:2206.02660

36. Morrison Philip J (1984) Bracket formulation for irreversible classical fields. Phys Lett A 100(8):423–427

37. Morrison PJ (1986) A paradigm for joined Hamiltonian and dissipative systems. Phys D Nonlinear Phenom 18(1–3):410–419

38. Grmela M (1984) Particle and bracket formulations of kinetic equations. Contemp Math 28:125–132

39. Grmela M (1984) Bracket formulation of dissipative fluid mechanics equations. Phys Lett A 102(8):355–358

40. Kaufman AN (1984) Dissipative Hamiltonian systems: a unifying principle. Phys Lett A 100(8):419–422

41. Öttinger HC, Grmela M (1997) Dynamics and thermodynamics of complex fluids. ii. illustrations of a general formalism. Phys Rev E 56(6):6633

42. Öttinger HC (2005) Beyond equilibrium thermodynamics. John Wiley & Sons, New York

43. Grmela M (2018) Generic guide to the multiscale dynamics and thermodynamics. J Phys Commun 2(3):032001

44. Grmela M, Klika V, Pavelka M (2019) Gradient and generic evolution towards reduced dynamics

45. Pavelka M, Klika V, Grmela M (2018) Multiscale thermodynamics. In : Multiscale thermo-dynamics. de Gruyter, Berlin

46. Quercus H, Badías A, González D, Chinesta F, Cueto E (2021) Structure-preserving neural networks. J Comput Phys 426:109950

47. Quercus H, Badias A, Gonzalez D, Chinesta F, Cueto E (2021) Deep learning of thermodynamics-aware reduced-order models from data. Comput Methods Appl Mech Eng 379:113763

48. Hernández Q, Badías A, Chinesta F, Cueto E (2022) Thermodynamics-informed graph neural networks. arXiv preprint arXiv:2203.01874

49. Moya B, Badias A, Gonzalez D, Chinesta F, Cueto E (2021) Physics perception in sloshing scenes with guaranteed thermodynamic consistency. arXiv preprint arXiv:2106.13301

50. Moya B, Badias A, Gonzalez D, Chinesta F, Cueto E (2022) Physics-informed reinforcement learning for perception and reasoning about fluids. arXiv preprint arXiv:2203.05775

51. Zhang Z, Shin Y, Em Karniadakis G (2022) Gfinns: generic formalism informed neural networks for deterministic and stochastic dynamical systems. Philos Trans R Soc A 380(2229):20210207

52. Lee K, Trask N, Stinis P (2021) Machine learning structure preserving brackets for forecasting irreversible processes. Adv Neural Inf Process Syst 34:5696–5707

53. Öttinger HC (2006) Nonequilibrium thermodynamics for open systems. Phys Rev E 73:036126

54. Badlyan AM, Maschke B, Beattie C, Mehrmann V (2018) Open physical systems: from generic to port-hamiltonian systems. arXiv preprint arXiv:1804.04064

55. Betsch P, Schiebl M (2018) Variational formulations for large strain thermo-elastodynamics based on the generic formalism. In: Proceedings of the 6th European conference on computational mechanics, Glasgow, UK, pp 11–15

56. Grmela M, Öttinger HC (1997) Dynamics and thermodynamics of complex fluids. i. development of a general formalism. Phys Rev E 56(6):6620

57. Romero I (2009) Thermodynamically consistent time-stepping algorithms for non-linear thermomechanical systems. Int J Numer Methods Eng 79(6):706–732