



Bergische Universität Wuppertal

Fakultät für Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational
Mathematics (IMACM)

Preprint BUW-IMACM 24/06

Luca Di Persio, Matthias Ehrhardt, and Sofia Rizzotto

Integrating Port-Hamiltonian Systems with Neural Networks: From Deterministic to Stochastic Frameworks

April 1, 2024

<http://www.imacm.uni-wuppertal.de>

Integrating Port-Hamiltonian Systems with Neural Networks: From Deterministic to Stochastic Frameworks

Luca Di Persio*, Matthias Ehrhardt†, Sofia Rizzotto‡

Department of Computer Science – College of Mathematics
University of Verona, Italy

IMACM, School of Mathematics and Natural Sciences,
University of Wuppertal, Germany

April 1, 2024

Abstract

This article presents an innovative approach to integrating port-Hamiltonian systems with neural network architectures, transitioning from deterministic to stochastic models. The study presents novel mathematical formulations and computational models that extend the understanding of dynamical systems under uncertainty and complex interactions. It emphasizes the significant progress in learning and predicting the dynamics of non-autonomous systems using port-Hamiltonian neural networks (pHNNs). It also explores the implications of stochastic neural networks in various dynamical systems.

AMS classification: 37N40, 37B52, 60G10, 60H10, 93C55

Keywords: Stochastic Port-Hamiltonian System, Port-Hamiltonian Neural Networks, Discrete stochastic Port-Hamiltonian system, Passivity, Interconnection

1 Introduction

This work presents an innovative integration of port-Hamiltonian systems (PHS) with neural network architectures, focusing in particular on the transition from deterministic models to those that incorporate stochastic elements. We first explore the advanced mathematical formulations and computational models to deepen the understanding of dynamical systems under uncertainty, including complex interactions and measurement errors. We first recall the definition of port-Hamiltonian systems as a synthesis of port modeling and geometric Hamiltonian dynamics, highlighting the use of the Dirac structure, which generalizes the concept of Poisson and pre-symplectic structures, to represent the energetic topology of a system. We then consider the stochastic extension of port-Hamiltonian systems to analyze uncertainties such as measurement errors and unknown environmental effects,

*luca.dipersio@univr.it

†Corresponding author, ehrhardt@uni-wuppertal.de

‡sofia.rizzotto@studenti.univr.it

making noise an intrinsic aspect of each port. For completeness, we also provide a detailed explanation of deterministic port-Hamiltonian systems using a coordinate-free geometric formulation.

The systems are described by generalized Poisson brackets and Hamiltonian functions, which are translated into equations involving Jacobian matrices and vector fields associated with Hamiltonian functions in local coordinates. We then consider the associated extension to implicit systems, which includes systems with algebraic constraints. These take into account the state space representing stored energy, vector spaces of flow and effort variables representing power ports, and the Dirac structure outlining the power-saving connections within the system, emphasizing a significant property of these systems: passivity. Special emphasis is placed on its relationship to the properties of the Dirac structure and its importance in ensuring stability in control systems. The transition from deterministic to stochastic *port-Hamiltonian neural networks* (pHNNs) is derived by incorporating noise directly into the structure via ports. This initiative represents a critical conceptual innovation and is motivated by real applications where uncertainties and the need to accommodate errors affect systems.

A port-Hamiltonian system combines aspects of port modeling with a coordinate-free geometric Hamiltonian dynamics, which describes the dynamics of the physical system and the interconnection structure through the equations of motion. The main feature of these systems lies in the geometric structure defining them, the so-called Dirac structure, which generalizes the concept of (pseudo) Poisson and pre-symplectic structures and represents the energetic topology of the system. This feature allows the framework to gain importance in modeling complex engineering systems with multiple energy exchange domains. The main focus of this paper is to explore an extension of port-Hamiltonian systems that incorporates random perturbations to account for uncertainties, measurement errors, and environmental interactions in dynamic systems. In particular, we will explore the innovation presented by Cordoni, Di Persio, and Muradore [2], where the uniqueness lies in considering noise as an intrinsic aspect of each port. This is motivated by the need to account for measurement errors, parameter estimation uncertainties, and the unknown effects of the environment on the system.

The paper is organized as follows. Section 1 establishes the basic concepts and explores different types of port-Hamiltonian systems, including discrete ones, and their application in neural networks. Section 1.1 lays the foundation by introducing essential concepts such as explicit input-state-output port-Hamiltonian systems, local coordinates, and implicit port-Hamiltonian systems. Next, Section 1.2 delves into a special type of port-Hamiltonian systems: discrete ones. It covers the generalized Dirac structure, a mathematical construct that underlies these systems, and how two such structures can be connected. Section 2 focuses on extending the concept of port-Hamiltonian systems to include random elements, making them suitable for modeling systems with uncertainties or randomness. We generalize the concept of Dirac structure to stochastic ones, to account for the stochastic nature introduced in this section. We also introduce the concept of a noise port, a special type of port through which random fluctuations or uncertainties enter the system. Section 3 extends the concept of passivity developed for deterministic systems to include the presence of randomness. Hereby, we explain strong and weak passivity and show how to determine whether a given stochastic port-Hamiltonian systems (SPHS) exhibits passivity. Section 4 focuses on applying and extending the concepts of stochastic PHS to various scenarios. First, Section 4.1 explores how to connect and analyze multiple SPHS systems. Section 4.2 adapts the continuous-time framework of SPHS to represent systems that evolve in discrete time steps. Section 4.3 discusses the application of SPHS to model the stochastic motion of agents, e.g. cars.

In Section 5 show successful applications of pHNNs in tasks such as simulating a damped mass-spring system and a chaotic Duffing system. These examples show that pHNNs can not only learn

system dynamics, but also recover complex behaviors such as chaotic trajectories with minimal data. Overall, pHNNs hold promise for several areas involving complex physical systems, including chemical interactions, robot motion control, and understanding general system dynamics without requiring precise details of the forces involved.

In the conclusions in Section 6, we summarize our work, which is based on the method of Colonijs and Grüne [14], which used neural networks for controller design. Our focus is on adapting this method to handle stochastic port-Hamiltonian systems with random fluctuations with the two key steps: 1) Modeling the system with noise: Introducing stochastic differential equations (SDEs) that include a term representing random noise (dW) to describe the system dynamics. 2) Design a Robust Controller: Modify the control law to account for the noise. This involves adding a new term ($\mu(x)$) to the original control law ($K(x)$) to counteract the stochastic disturbances and ensure that the system remains stable and performs well.

Finally, in A we provide a discussion of stochastic neural networks (SNNs), which are inspired by the biological brain. Unlike traditional deterministic networks, SNNs incorporate randomness to improve training and avoid overfitting.

1.1 Basic Concepts

The description given by Cordonì, Di Persio and Muradore [2] of an input-state-output deterministic port-Hamiltonian system can be done by considering a geometric coordinate-free formulation in terms of Poisson brackets, i.e. the following system

$$\begin{cases} \dot{x} &= X_H(x) + \sum_{i=1}^m u_i X_{H_{g_i}}(x), \\ y_i &= \{H, H_{g_i}\}, \end{cases} \quad (1)$$

is called a (*explicit*) *input-state-output port-Hamiltonian system* (PHS) on a Poisson manifold $(X, \{\cdot, \cdot\})$ with a Hamiltonian function $H \in C^\infty(\mathcal{X})$, $x \in \mathbb{R}^n$, the i -th input $u_i \in U$, the i -th output $y_i \in U^*$ and the Hamiltonian vector field $X_{H_{g_i}}$ associated with the Hamiltonian H_{g_i} . In local coordinates, the previous system reads

$$\begin{cases} \dot{x} &= J(x)\partial_x H + \sum_{i=1}^m u_i g_i(x), \\ y_i &= g_i^\top(x)\partial_x H. \end{cases} \quad (2)$$

Moreover, considering $X_H^L(\cdot) := [\cdot, H]_L$, we can define the (*explicit*) *input-state-output port-Hamiltonian system with dissipation* as

$$\begin{cases} \dot{x} &= X_H^L(x) + \sum_{i=1}^m u_i H_{g_i}(x), \\ y_i &= [H, H_{g_i}], \end{cases} \quad (3)$$

and in local coordinates this reads

$$\begin{cases} \dot{x} &= (J(x) - R(x))\partial_x H(x) + \sum_{i=1}^m u_i g_i(x), \\ y_i &= g_i^\top(x)\partial_x H(x), \end{cases} \quad (4)$$

where $R(x) := (g^R(x))^\top \tilde{R}(x)g^R(x)$.

Let's consider a physical system consisting of elements that store energy, a set of elements that dissipate energy, and a set of power ports interconnected by power-preserving links that can only transfer energy, not produce it. We can describe such a system by extending the framework of port-Hamiltonian systems to the context of implicit systems, i.e., systems with algebraic constraints. Given a state space \mathcal{X} (a smooth manifold whose elements represent the energy stored in the system),

a vector space of flow variables \mathcal{V} and its dual space of effort variables \mathcal{V}^* (representing the power ports), a geometric Dirac structure \mathcal{D} and a Hamiltonian function \mathcal{H} representing the total energy of the system in a given state, we can define an *implicit port-Hamiltonian system* corresponding to $(\mathcal{X}, \mathcal{V}, \mathcal{D}, \mathcal{H})$ as

$$v = -\dot{x} \quad \text{and} \quad v^* = \frac{\partial \mathcal{H}}{\partial x}(x), \quad (5)$$

which means that the system is defined by

$$\left(-\dot{x}, \frac{\partial \mathcal{H}}{\partial x}(x), f, e\right) \in \mathcal{D}(x). \quad (6)$$

Note that the above Dirac structure outlines the behavior of the internal connection, and its main property is that the power-conserving combination of Dirac structures remains a Dirac structure. This implies that any power-conserving connection of port-Hamiltonian systems is also a port-Hamiltonian system, where the Dirac structure is the composition of the Dirac structures of its elementary parts, and the Hamiltonian is the sum of the Hamiltonians.

Another important property of a system is passivity, which means that the energy input to the system is always greater than or equal to (lossless case) the energy output from the system. It is crucial for ensuring stability in control systems. Passivity is a consequence of the properties of the Dirac structure and the energy-dissipation relation for port-Hamiltonian systems.

1.2 Discrete Systems

So far, we have described port-Hamiltonian systems in continuous time. However, it is essential to investigate whether these properties persist when we discretize time. As suggested by Viswanath, Clemente-Gallardo and van der Schaft [4], the discretization of a Hamiltonian system can be done while preserving energy conservation. Discrete systems can be obtained by discretizing a continuous system or by modelling systems directly at the discrete level. To obtain the latter, we can use Poisson brackets, since they satisfy the properties of skew symmetry, bilinearity, and a modified Leibniz rule, which are necessary to preserve the structure of the Hamiltonian dynamics in the discrete setting. First, we need to define the discrete Dirac structure. For this purpose we denote with $\mathfrak{X}(A)$ a space of discrete vector fields, with $\Lambda^1(A)$ the space of discrete 1-forms. Then a *generalized Dirac structure* on an n -dimensional discrete manifold is an n -dimensional linear subspace $\mathcal{D} \subset \mathfrak{X}(A) \times \Lambda^1(A)$ such that $\mathcal{D} = \mathcal{D}^\top$ with

$$\mathcal{D}^\top = \{(Y, \beta) \in \mathfrak{X}(A) \times \Lambda^1(A) \quad \text{where} \quad \langle \alpha, X \rangle + \langle \beta, Y \rangle = 0, \quad \forall (X, \alpha) \in \mathcal{D}\},$$

where $\langle \cdot, \cdot \rangle$ is the pairing between \mathbb{F}^n and \mathbb{F}^{n*} . Then, assume to have the situation illustrated in Figure 1, denoting with \mathcal{F}_i the space of flow and with \mathcal{F}_i^* the space of effort of the Dirac structure \mathcal{D}_i with $i = A, B$, we can define the *interconnection between the two Dirac structures* \mathcal{D}_A and \mathcal{D}_B as

$$\begin{aligned} \mathcal{D}_A \circ \mathcal{D}_B := \{ & (f_1, e_1, f_2, e_2) \in \mathcal{F}_1 \times \mathcal{F}_1^* \times \mathcal{F}_2 \times \mathcal{F}_2^* \quad \text{such that} \\ & \exists (f, e) \in \mathcal{F} \times \mathcal{F}^* \quad \text{with} \quad (f_1, e_1, f, e) \in \mathcal{D}_A \quad \text{and} \quad (-f, e, f_2, e_2) \in \mathcal{D}_B \}. \end{aligned} \quad (7)$$

Then $\mathcal{D}_A \circ \mathcal{D}_B$ is a Dirac structure.

Definition 1 (Implicit discrete port-Hamiltonian system [4]). *Let $H : \mathcal{Z} \rightarrow \mathbb{F}$ be a discrete Hamiltonian, \mathcal{F}_P be the space of external flows f , $\mathcal{E}_P = \mathcal{F}_P^*$ the space of external effort e , \mathcal{D} be the Dirac structure which depends only on the coordinate z , then the implicit discrete port-Hamiltonian system is defined as*

$$\left(-\frac{\Delta z}{\Delta t}, f, \partial_z H(z), e\right) \in \mathcal{D}(z). \quad (8)$$

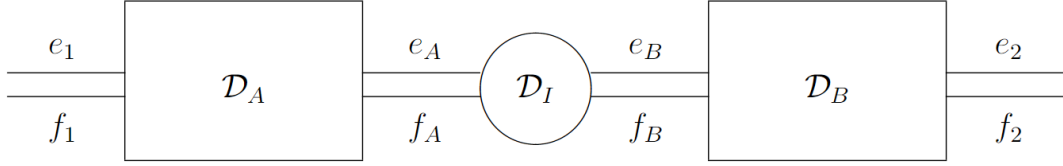


Figure 1: Interconnection of two port-Hamiltonian systems.

There are also some cases where the system consists of an interconnection of smooth and discrete port-Hamiltonian systems, which can be achieved (as proposed by Viswanath, Clemente-Gallardo and van der Schaft [6]) by transforming the problem of energy-conserving connection on sampling instances, into an energy-conserving interconnection with an external flow source. In this way, we ensure energy conservation over sampling intervals. Moreover, the resulting interconnected system is again passive and port-Hamiltonian. Thus, the integration of discrete computational models with physical systems does not compromise the stability and performance characteristics of passive systems. However, energy conservation does not hold, so the idea of Kotyczka and Lefevre [5] is to introduce a discrete-time Dirac structure and a discrete-time port-Hamiltonian system. This method approximates the continuous-time structural energy balance and uses symplectic numerical time integration by collocation methods.

2 Stochastic port-Hamiltonian systems

Consider a complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P})$ and denote by δZ the Stratonovich integration and by dZ the Itô integration along the semimartingale Z . As mentioned above, the stochastic port-Hamiltonian formulation allows for a more extensive source of randomness by treating each element of the system as a semimartingale. The adoption of the Stratonovich calculus is due to the geometric nature of the Dirac structure, but we can translate it into the corresponding Itô formulation to exploit its probabilistic properties. Consider the following system

$$\begin{cases} \delta X_t &= (J(X_t) - R(X_t)) \partial_x H(X_t) \delta Z_t + g(X_t) u \delta Z_t^g + \xi(X_t) \delta Z_t^N, \\ y_t &= g^\top(X_t) \partial_x H(X_t), \end{cases} \quad (9)$$

where $R(x) := (g^R(x))^\top \tilde{R}(x) g^R(x)$, W is a Brownian motion, and Z , Z^g , and Z^N are semimartingales. Then (9) describes the stochastic port-Hamiltonian system in local coordinates.

By generalizing the notion of Dirac structure, it is possible to extend the previously described framework by incorporating scenarios where noise is introduced into the system as a stochastic external random field and as a random perturbation of any port connected to the system.

Definition 2 (Orthogonal complement [2]). *Given a manifold \mathcal{X} , $I \subset \mathbb{R}_+$, a bundle $\mathcal{D} \subset T\mathcal{X} \oplus T^*\mathcal{X}$, a differential 1-form σ on \mathcal{X} and an integral curve $X : I \rightarrow \mathcal{X}$ of a Stratonovich vector field δX_t , the orthogonal complement of \mathcal{D} is*

$$\mathcal{D}^\perp = \{(\delta X_t, \sigma) \in T\mathcal{X} \oplus T^*\mathcal{X} : \int_0^t \langle \sigma, \delta \bar{X}_s \rangle + \int_0^t \langle \bar{\sigma}, \delta X_s \rangle = 0, \forall (\delta \bar{X}_t, \bar{\sigma}) \in \mathcal{D}, t \in I\} \quad (10)$$

Definition 3 (Generalized Dirac structure [2]). *With the same notation as above, we call generalized stochastic Dirac structure a smooth vector subbundle $\mathcal{D} \subset T\mathcal{X} \oplus T^*\mathcal{X}$ such that $\mathcal{D} = \mathcal{D}^\perp$.*

Definition 4 (Implicit generalized stochastic PHS [2]). *Let $H : \mathcal{X} \rightarrow \mathbb{R}$ be a Hamiltonian function, Z a semimartingale perturbing the system, then an implicit generalized stochastic port-Hamiltonian system on \mathcal{X} is a 4-tuple $(\mathcal{X}, Z, \mathcal{D}, H)$ such that*

$$(\delta X_t, \mathbf{d}H(X_t)) \in \mathcal{D}(X_t) \quad \forall t \in I. \quad (11)$$

Including a resistive element and an external element control then an implicit generalized port-Hamiltonian system with resistive structure \mathcal{R} is a 5-tuple $(\mathcal{X}, Z, \mathcal{F}, \mathcal{D}, H)$ such that

$$(-\delta X_t, \mathbf{d}H, \delta f_t^R, e_t^R, \delta f_t^C, e_t^C) \in \mathcal{D}(\mathcal{X})_\square \quad \text{with} \quad (\delta f_t^R, e_t^R) \in \mathcal{R}(\mathcal{X})_t.$$

Proposition 1. [2] *Implicit port-Hamiltonian systems satisfy energy conservation property that is*

$$H(X_t) - H(X_0) = \int_0^t \langle \mathbf{d}H, \delta X_s \rangle \quad (12)$$

which can be written in short notation as

$$\delta H(X_t) = \langle \mathbf{d}H, \delta X_t \rangle. \quad (13)$$

The energy balance is

$$H(X_t) - H(X_0) = \int_0^t \langle e_s^R, \delta f_s^R \rangle + \int_0^t \langle e_s^C, \delta f_s^C \rangle \leq \int_0^t \langle e_s^C, \delta f_s^C \rangle. \quad (14)$$

However, the condition $\int_0^t \langle e_s^R, \delta f_s^R \rangle \leq 0$ imposed on the resistive port is too strong since it is difficult for it to happen in practice, so the idea is to introduce the following weaker definition of the resistive relation $\mathcal{R}_W \subset \mathcal{F}_{Z_R} \times \mathcal{E}_R$:

$$\mathbb{E} \int_0^t \langle e_s^R, \delta f_s^R \rangle \leq 0. \quad (15)$$

Consequently, the mean power balance requires that the energy be conserved and dissipated in mean value, i.e. it reads

$$\mathbb{E}(H(X_t) - H(X_0)) = \mathbb{E} \int_0^t \langle e_s^R, \delta f_s^R \rangle + \mathbb{E} \int_0^t \langle e_s^C, \delta f_s^C \rangle \leq \mathbb{E} \int_0^t \langle e_s^C, \delta f_s^C \rangle. \quad (16)$$

It is possible to further generalize the Hamiltonian by introducing an external perturbation of the system, i.e. a new type of port called *noise port* perturbed by the semimartingale Z_N (see Figure 2). Thus, in this case the *implicit generalized stochastic port-Hamiltonian system with resistive structure* is a 5-tuple $(\mathcal{X}, Z, \mathcal{F}, \mathcal{D}, H)$ such that

$$(-\delta X_t, \mathbf{d}H, \delta f_t^R, e_t^R, \delta f_t^C, e_t^C, \delta f_t^N, e_t^N) \in \mathcal{D}(X_t) \quad (17)$$

and the weak energy balance is given by

$$\mathbb{E}H(X_t) - \mathbb{E}H(X_0) \leq \mathbb{E} \int_0^t \langle e_s^N, \delta f_s^N \rangle + \mathbb{E} \int_0^t \langle e_s^C, \delta f_s^C \rangle. \quad (18)$$

Furthermore, we can reformulate the stochastic port-Hamiltonian system in Itô form by defining the following alternative representation of the Dirac structure.

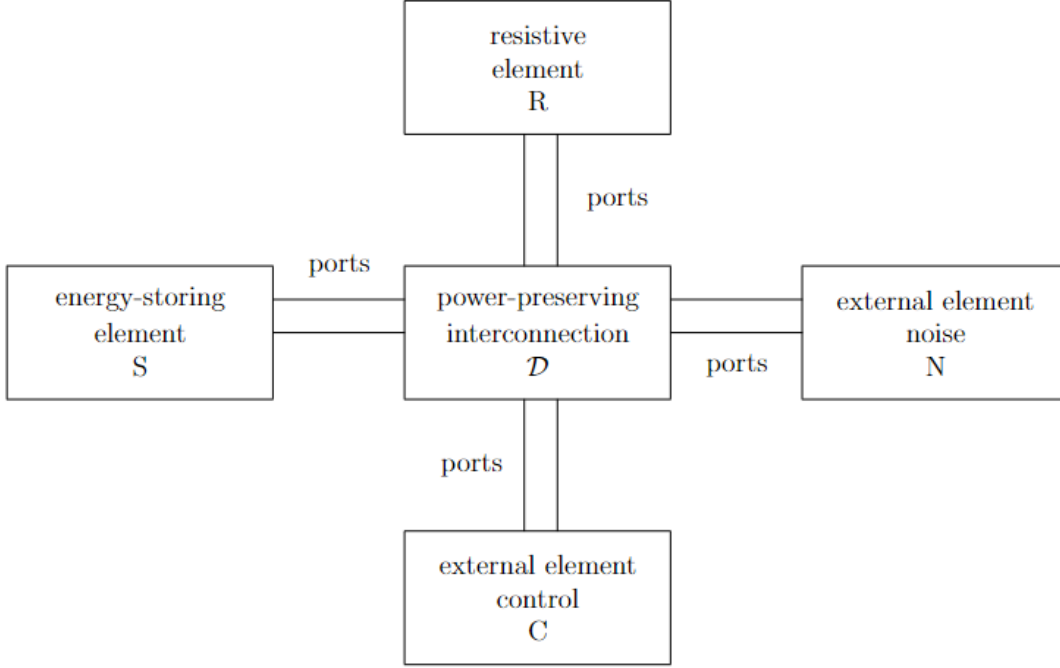


Figure 2: Schematic representation of a general implicit port-Hamiltonian system.

Definition 5 (Dirac structure [2]). Let $\mathcal{F} := \mathcal{F}_{Z_R} \times \mathcal{F}_{Z_C} \times \mathcal{F}_{Z_N}$ be the space of flows δf , $\mathcal{E} = \mathcal{F}^*$ be the dual space of efforts e , $G_\theta : \mathcal{F}_{Z_\theta} \rightarrow \mathcal{F}_{Z_\theta}$ be a function such that $\langle e_t^S, G_\theta \delta f_t^\theta \rangle = \langle G_\theta^* e_t^S, \delta f_t^\theta \rangle$ and J be a matrix such that $J = -J^\top$. Then the Dirac structure \mathcal{D} can be defined as

$$\begin{aligned} \mathcal{D} := \{ & (\delta f_t^S, \delta f_t^R, \delta f_t^C, \delta f_t^N, e_t^S, e_t^R, e_t^C, e_t^N) \in \mathcal{F} \times \mathcal{E} : \\ & \delta f_t^S = -J e_t^S \delta Z_t - G_R \delta f_t^R - G_C \delta f_t^C - G_N \delta f_t^N, \\ & e_t^R = G_R^* e_t^S, \quad e_t^C = G_C^* e_t^S, \quad e_t^N = G^* e_t^S \}. \end{aligned} \quad (19)$$

Consider the special case

$$\begin{aligned} \delta f_t^R &= -\tilde{R} e_t^R \delta Z_t, \quad \delta f_t^N = \xi_t \delta Z_t^N, \quad \delta f_t^C = u_t \delta Z_t^C \\ \text{with } \mathbb{E} \int_0^t \langle e_s^R, \tilde{R} e_s^R \delta Z_s \rangle - \mathbb{E} \int_0^t \langle e_s^N, f_s^N \delta Z_s^N \rangle &\geq 0, \end{aligned} \quad (20)$$

where the reason for the minus sign in front of \tilde{R} is that we want it to be the incoming power regarding the interconnection (as in [1]).

Definition 6 (Stochastic input-output PHS with stochastic Dirac structure [2]). Using the same notation as above, if $\mathbf{Z} = (Z, Z^R, Z^C, Z^N)$ is a semimartingale and $H : \mathcal{X} \rightarrow \mathbb{R}$ is a Hamiltonian function, then the stochastic input-output port Hamiltonian system with stochastic Dirac structure is given by

$$\begin{cases} \delta X_t &= -J dH(X_t) \delta Z_t + G_R \tilde{R} e_t^R \delta Z_t - G_C u_t \delta Z_t^C - G_N \xi_t \delta Z_t^N, \\ e_t^N &= G_N^* dH(X_t), \\ e_t^C &= G_C^* dH(X_t) \end{cases} \quad (21)$$

and by taking $\tilde{J} = -J$ and $e_t^R = G_R^* \mathbf{d}H(X_t)$, the system 21 becomes

$$\begin{cases} \delta X_t &= (\tilde{J} + G_R \tilde{R} G_R^*) \mathbf{d}H(X_t) \delta Z_t - G_C u_t \delta Z_t^C - G_N \xi_t \delta Z_t^N, \\ e_t^N &= G_N^* \mathbf{d}H(X_t), \\ e_t^C &= G_C^* \mathbf{d}H(X_t). \end{cases} \quad (22)$$

Theorem 1. [2] If X is a solution of the equation (22) and Z, Z^N, Z^C are such that

$$\langle Z, Z^C \rangle_t = \langle Z, Z^N \rangle_t = \langle Z, Z^C \rangle_t = 0$$

where $\langle \cdot, \cdot \rangle_t$ is the quadratic covariation at time t , then X can be equivalently rewritten in Itô terms as

$$\begin{aligned} dX_t &= V^S(X_t) dZ_t + \mathcal{L}_{V^S} V^S(X_t) d\langle Z, Z \rangle_t + \\ &\quad - \sum_{i=1}^{n^N} V_i^N(X_t) dZ_t^N - \frac{1}{2} \sum_{i,j=1}^{n^N} \mathcal{L}_{C_j^N} V_i^N(X_t) d\langle Z^{N:i}, Z_t^{N:j} \rangle + \\ &\quad - \sum_{i=1}^{n^C} V_i^C(X_t) u_t^i dZ_t^{C:i} - \frac{1}{2} \sum_{i,j=1}^{n^C} \mathcal{L}_{V_j^C} V_i^C(X_t) u_t d\langle Z^{C:i}, Z^{C:j} \rangle_t, \end{aligned} \quad (23)$$

where \mathcal{L} is the Lie derivative and V^α , $\alpha = S, N, C$ are defined as

$$(\tilde{J} + G_R \tilde{R} G_R^*) \mathbf{d}H = V^S, \quad G_N \xi_t = \sum_{i=1}^{n^N} V_i^M, \quad G_C = \sum_{i=1}^{n^C} V_i^C. \quad (24)$$

3 Passivity in Stochastic Systems

Extending the notion of passivity to the stochastic case is challenging because the presence of noise affects the energy of the system. The standard requirement that the structure matrix R be symmetric and positive semidefinite is no longer sufficient due to the presence of the semimartingale Z , which makes the Stochastic port-Hamiltonian System (SPHS) non-dissipative, so it is necessary to impose specific conditions on the noise to ensure losslessness and passivity.

Proposition 2. [2] If X is the solution of an explicit I-S-O stochastic PHS with dissipation, i.e.

$$\begin{cases} \delta X_t &= X_H^L(X_t) \delta Z_t + u X_{H_g}^L(X_t) \delta Z_t^g + X_{H_N}^L(X_t) \delta Z_t^N, \\ y_t &= [H, H_g]_L, \end{cases} \quad (25)$$

with

$$X_H^L(\cdot) := [\cdot, H]_L, \quad X_{H_g}^L(\cdot) := [\cdot, H_g]_L, \quad X_{H_N}^L(\cdot) := [\cdot, H_N]_L,$$

then for all $\varphi \in C^\infty(\mathcal{X})$ it holds

$$\begin{cases} \delta \varphi(X_t) &= [\varphi, H]_L(X_t) \delta Z_t + u [\varphi, H_g]_L(X_t) \delta Z_t^g + [\varphi, H_N]_L(X_t) \delta Z_t^N, \\ y_t &= [H, H_g]_L. \end{cases} \quad (26)$$

The equation changes when there is no external noise and the semimartingale perturbing the control is deterministic.

Definition 7 (Strong and weak passivity [2]). *If $H \in C^\infty(\mathcal{X})$ is the total energy of the explicit I-S-O stochastic PHS with dissipation, then it is strongly passive if for all $t \geq 0$ it holds*

$$H(X_t) \leq H(X_0) + \int_0^t u^\top(s)y(s) \delta Z_s^C, \quad (27)$$

or weakly passive if for all $t \geq 0$ it holds

$$\mathbb{E}H(X_t) \leq \mathbb{E}H(X_0) + \mathbb{E} \int_0^t u^\top(s)y(s) \delta Z_s^C. \quad (28)$$

Assuming $\xi = 0$, the energy conservation relation of the system

$$\begin{cases} \delta X_t &= (J(X_t) - R(X_t)) \partial_x H(X_t) \delta Z_t + g(X_t)u \delta Z_t^C + \xi(X_t) \delta Z_t^N, \\ e^C &= g^\top(X_t) \partial_x^H(X_t), \end{cases} \quad (29)$$

where $J = J^\top$, is given by

$$\begin{aligned} H(X_t) - H(X_0) &= \int_0^t \langle \mathbf{d}H, \delta X_s \rangle = \int_0^t \langle e_s^R, \delta f_s^R \rangle + \int_0^t \langle e_s^C, \delta f_s^C \rangle \\ &= \int_0^t \langle \partial_x H(X_s), R(X_s) e^R \delta Z_s \rangle + \int_0^t \langle y, u \delta Z_s^C \rangle \\ &= \int_0^t \langle \partial_x H(X_s), -R(X_s) \partial_x H(X_s) \delta Z_s \rangle + \int_0^t y^\top u \delta Z_s^C \\ &= - \int_0^t \partial_x^\top H(X_s) R(X_s) \partial_x H(X_s) \delta Z_s + \int_0^t y^\top u \delta Z_s^C. \end{aligned} \quad (30)$$

We can see that even if R is strictly positive, we cannot infer that the strongly passive condition holds; in fact, we should also require that

$$\int_0^t \partial_x^\top H(X_s) R(X_s) \partial_x H(X_s) \delta Z_s \geq 0. \quad (31)$$

However, the latter condition is usually difficult to satisfy in a real system, so it makes more sense to require the weaker condition.

$$\mathbb{E} \int_0^t \partial_x^\top H(X_s) R(X_s) \partial_x H(X_s) \delta Z_s \geq 0, \quad (32)$$

which guarantee the weak passivity property. However, computing the expectation of a Stratonovich integral is often challenging. The natural solution to this problem is to transform the equation (28) into Itô terms. This transformation allows us to take advantage of the favourable probabilistic properties of the Itô integral. To achieve this, we can apply the Theorem 1 to the equation (28).

4 Application and Generalization

4.1 Interconnection of Multiple SPHS

An essential property of port-Hamiltonian systems is their interconnectivity, which allows complex systems to be viewed as compositions of simpler parts. This interconnectivity can be analysed regarding the components and how they are interconnected. In particular, through the composition of

Dirac structures, the power-preserving interconnection of port-Hamiltonian systems defines another port-Hamiltonian system. The Hamiltonian of the interconnected port-Hamiltonian system is the sum of the Hamiltonians of its components, and the energy-dissipation relation is the union of the energy-dissipation relations of the subsystems.

The following proposition discusses the connection of multiple SPHS, defining a new system with interconnected Dirac structures and combined Hamiltonians.

Proposition 3. *Suppose we have N stochastic port-Hamiltonian systems with state space \mathcal{X}_i , Hamiltonian H_i , flow-effort space $\mathcal{F}_i \times \mathcal{E}_i$ and perturbation \mathbf{Z}^i for $i = 1, \dots, N$. Assuming that they are connected by \mathcal{D}_I (see Figure 3), then their interconnection defines a stochastic port-Hamiltonian system with Dirac structure $\mathcal{D} \circ \mathcal{D}_I$ and Hamiltonian $H := \sum_{i=1}^N H_i$.*

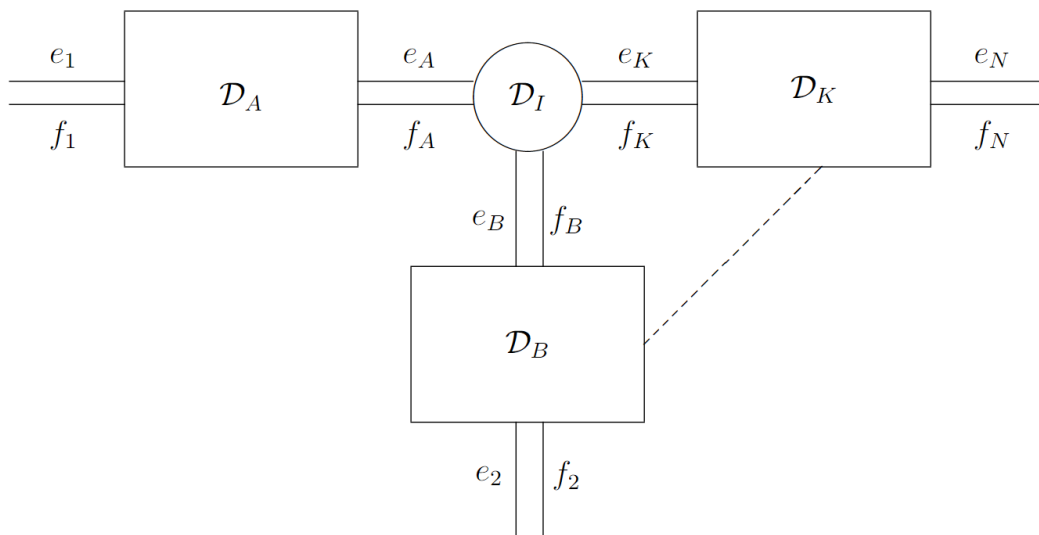


Figure 3: Interconnection of N implicit port-Hamiltonian systems.

4.2 Discrete Stochastic PHS

Consider the continuous stochastic port-Hamiltonian system

$$\begin{cases} dX_t &= ((J - R)\partial_x H(X_t) + g(X_t)u_t) dt + \xi(X_t) \delta W_t, \\ y_t &= g^\top(X_t)\partial_x H(X_t), \\ z_t &= \xi^\top(X_t)\partial_x H(X_t), \end{cases} \quad (33)$$

where $J = -J^\top$, R positive semidefinite, g represents control port, H is the Hamiltonian, $u \in U$ is the control input, $y \in Y$ is the output of the system, ξ is a matrix, z is the associated effort to δW_t and W is a standard Brownian motion adapted to the reference filtration $(\mathcal{F}_t)_{t \geq 0}$. Then we can introduce the discretization

$$\dot{X}(t_0^k + \tau h) = -f(t_0^k + \tau h) = -\sum_{j=1}^s f_j^k l_j(\tau) \quad (34)$$

with

$$\dot{X}(t_i^k) := -f_i^k, \quad l_i = \prod_{j=1}^s \frac{\tau - c_j}{c_i - c_j}, \quad \tau \in [0, 1],$$

where l_i is the i^{th} Lagrange interpolation polynomial of order s and τ is the normalized time parameterizing the sampling intervals. Thus, we can generalize the continuous SPHS to a discrete form, preserving the structure of the Hamiltonian and the control inputs, as follows:

Definition 8 (Discrete stochastic port-Hamiltonian system [5]). *A discrete stochastic port-Hamiltonian system can be written as*

$$\begin{cases} X(t_0^k + c_i h) &= x_0^k - h \sum_{i=1}^s a_{ij} f_j^k, \\ X(t_0^k + h) &= x_0^k - h \sum_{i=1}^s a_j f_i^k, \\ -a_{ij} f^k &= (J_j^k - R_j^k) a_{ij} e_j^k + a_{ij} g_j^k u_j^k + b_{ij} \xi_j^k \Delta W, \end{cases} \quad (35)$$

where ΔW is a truncated centred Gaussian random variable with variance h , $a_{ij} = \int_0^{c_i} l_j(\sigma) d\sigma$, $a_j = \int_0^1 l_j(\sigma) d\sigma$ and $M = M^\top$ [5].

Note that in the discrete case, the system is *passive* if it holds

$$\mathbb{E}[\Delta H^k] \leq h \mathbb{E}[(y^k)^\top u^k]. \quad (36)$$

4.3 Stochastic Motion Model of Agents

An application to a stochastic motion model of agents is presented by Ehrhardt, Kruse, and Tordeux [12], who analyze the case where positions and velocities of agents are modeled in a ring structure. The initial positions and velocities are set, and the system is governed by differential equations involving the velocities and positions of neighboring agents [12]. The dynamic equations of the agents are given by:

$$\begin{cases} dQ_n(t) &= (p_{n+1}(t) - p_n(t)) dt, \\ dp_n(t) &= (U'(Q_n(t)) - U'(Q_{n-q}(t))) dt + \\ &\quad + \beta(p_{n+1}(t) - 2p_n(t) + p_{n-1}(t)) dt + \sigma dW_n(t) \end{cases} \quad (37)$$

with $Q(0) = Q_0 \in [0, +\infty)^N$ the initial distance, $p(0) = p_0$ the initial velocity, $\beta \in (0, +\infty)$ a dissipation rate, $\sigma \in \mathbb{R}$ the noise volatility, U' the derivative of a convex potential $U \in C^1(\mathbb{R}, [0, +\infty))$ and $W = (W_n)_{n=1}^N: [0, +\infty) \times \Omega \rightarrow \mathbb{R}^N$ an N -dimensional standard Brownian motion defined on $(\Omega, \mathcal{F}, \mathbb{P})$. These equations represent the acceleration of the n^{th} agent depending on the velocities of its neighbors and the stochastic perturbations of the Brownian motion [13].

The motion of the agents is further formulated using a stochastic port-Hamiltonian framework

$$dZ(t) = (J - R)\nabla H(Z(t)) dt + G dW(t),$$

where $Z(t) = (Q(t), p(t))^\top \in \mathbb{R}^{2N}$, $t \in [0, +\infty)$, J and R are defined as skew-symmetric and symmetric positive semidefinite matrices, respectively. This formulation allows the application of Hamiltonian dynamics to model agents' behaviour under stochastic influences. In particular, the Hamiltonian is independent of Q and U , and its expectation could increase with time. Moreover,

describing the limiting behavior of these stochastic systems is challenging, so the authors [12] focus on the specific scenario where the quadratic function characterizes the potential

$$U(x) = \frac{(\alpha x)^2}{2} \quad x \in \mathbb{R}, \quad \alpha \in (0, \infty), \quad (38)$$

in which the process reads

$$dZ(t) = BZ(t) dt + G dW(t), \quad Z(0) = (Q_0, p_0)^\top, \quad (39)$$

where B is defined such that $BZ(t) = (J - R)\nabla H(Z(t))$. In this case, the resulting process converges for $t \rightarrow \infty$ in distribution to a normal distribution with known expectation and covariance matrix.

5 Results and Discussion

To introduce the port-Hamiltonian formalism into the neural network architecture, we can analyze the behavior of port-Hamiltonian Neural Networks (pHNN) on different tasks such as the mass-spring system with the damped term, external force and the Duffing equations. In particular, the port-Hamiltonian Neural Network (pHNN) is a significant advance in learning and predicting the dynamics of non-autonomous systems. Many real-world dynamical systems involve time-dependent forces and energy dissipation, which pose a challenge for learning. Desai et al. [3] evaluate the pHNN on several tasks, including a mass-spring system with damping and an external force, and a Duffing system. In particular, the pHNN can visually recover the Poincaré section of a chaotically driven system, highlighting its potential to identify and understand chaotic trajectories with minimal data. Its applicability extends to complex nonlinear forced and damped physical systems, encouraging applications in areas such as chemical bonding forces, robotic motion, and controlled dynamics without explicit knowledge of force and damping.

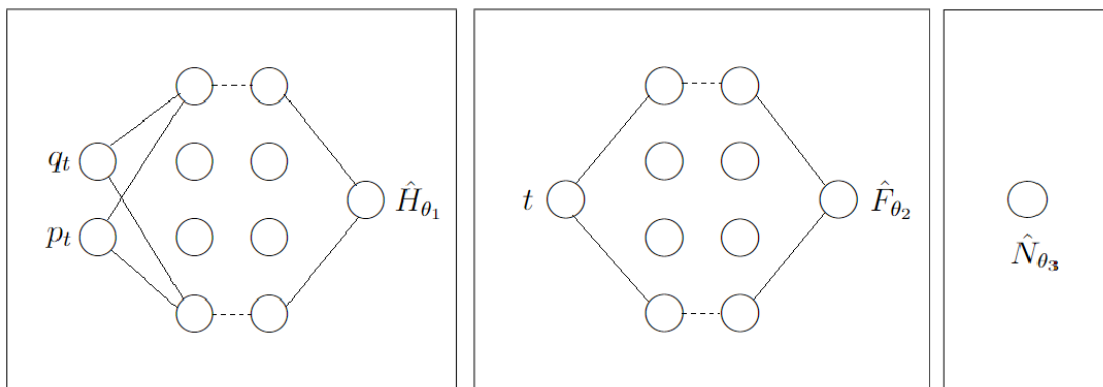


Figure 4: Schematic representation of a port-Hamiltonian neural network (pHNN).

As we can see in Figure 4, the main idea is to use port Hamiltonian theory to explicitly learn the force F_{θ_2} , the damping term N_{θ_3} and the Hamiltonian \mathcal{H}_{θ_1} . This should allow us to predict the time derivatives

$$\begin{bmatrix} \hat{q}_t \\ \hat{p}_t \end{bmatrix} = \begin{bmatrix} \frac{d\hat{H}_{\theta_1}}{dp_t} \\ -\frac{d\hat{H}_{\theta_1}}{dq_t} + \hat{N}_{\theta_3} \frac{d\hat{H}_{\theta_1}}{dp_t} + \hat{F}_{\theta_2} \end{bmatrix}. \quad (40)$$

Furthermore, knowing $[\dot{q}, \dot{p}]$ from the data, we can compute the loss function [15]

$$\mathcal{L} = \underbrace{\|\hat{q}_t - \dot{q}_t\|_2^2 + \|\hat{p}_t - \dot{p}_t\|_2^2}_{\text{first part}} + \underbrace{\lambda_F \|\hat{F}_{\theta_2}\|_1 + \lambda_N \|\hat{N}_{\theta_3}\|_1}_{\text{second part}}. \quad (41)$$

In particular, the procedure for implementing the port-Hamiltonian neural network is as follows:

- Obtain ground truth state variable data $[q, p, t]$ and time derivatives $[\dot{q}, \dot{p}]$ from trajectories of a given system;
- Provide state variable information to pHNN, which learns a Hamiltonian, force, and damping term to predict $[\hat{q}, \hat{p}]$;
- Optimize pHNN by minimizing the loss function of equation (41);
- Train the system;
- Use the pHNN in a scientific integrator to evolve a set of random initial conditions in the test set;
- Estimate the goodness of the performance by computing the mean square error of the predicted state variables.

To understand the power of this method, we again compare different neural network models: the baseline neural network, the Hamiltonian neural network, the time-dependent Hamiltonian neural network, and the port-Hamiltonian neural network, against ground truth data [3].

In a simple mass-spring system, characterized by mass m and spring constant k , the time-independent Hamiltonian is given by

$$H = \frac{1}{2}kq^2 + \frac{p^2}{2m}.$$

After training the system and testing it over 25 random initial conditions, we find that both the Hamiltonian Neural Network (HNN) and the port-Hamiltonian Neural Network (pHNN) produce similar results in terms of state and energy mean squared error (MSE). In particular, pHNN successfully learns to predict force and damping over time, even when the ground truth observations are null. However, the prediction error remains small, approximately on the order of 10^{-5} and 10^{-8} , respectively. Furthermore, comparing the state and energy MSE among the listed methods, it is evident that HNN achieves the lowest MSE, followed by pHNN, time-dependent Hamiltonian neural network (TDHNN), and the baseline neural network (bNN).

In the damped mass-spring system, we introduce a damping term that gradually decreases the initial energy of the system over time. Denoted by δ , the damping coefficient, the system is described by the equation

$$\dot{q} = -\delta q.$$

Furthermore, based on the definition $p = m\dot{q}$, we have:

$$\dot{q} = \frac{1}{m}p.$$

By integration, we get $q^\top q - \delta q^\top \dot{q}$, which, when combined with a kinetic energy term, gives the Hamiltonian:

$$H = \frac{p^\top p}{2m} + \frac{q^\top q}{2} - \delta q^\top \dot{q}.$$

However, this formulation violates the previous equation, since $\frac{\partial H}{\partial p} = \dot{q} - \delta m q^{-1} \dot{q} \neq \dot{q}$. Consequently, including this term violates energy conservation, making it impossible to derive a scalar Hamiltonian for such a system. After training and testing this system, it is observed that both the Hamiltonian Neural Network (HNN) and the Time Dependent Hamiltonian Neural Network (TDHNN) fail to capture the dynamics of the system accurately. However, the baseline neural network (bNN) and the *port-Hamiltonian neural network* (pHNN) show better performance, effectively recovering the system dynamics and minimizing the state and energy mean square error (MSE). Moreover, in this scenario, the pHNN converges to forcing and damping terms consistent with the ground truth. In the forced mass-spring system, we introduce a forced time-dependent force that governs the undamped system. A Hamiltonian describing this forced mass-spring system has the form

$$H = \frac{1}{2}kq^2 + \frac{p^2}{2m} - qF_0 \sin(\omega t).$$

Here, F_0 is the force amplitude and ω is the frequency of the external force term. After training and testing, both the Hamiltonian neural network (HNN) and the time-dependent Hamiltonian neural network (TDHNN) have difficulty learning the dynamics due to the lack of explicit time dependence in the Hamiltonian. Conversely, the port-Hamiltonian neural network (pHNN) exhibits the lowest state, and energy mean squared error (MSE), indicating better performance. In addition, the baseline neural network (bNN) has a relatively low MSE. However, when we consider a more complex force term, such as

$$H = \frac{1}{2}kq^2 + \frac{p^2}{2m} - qF_0 \sin(\omega t) \sin(2\omega t),$$

where the complexity of the force significantly affects the performance of bNN. In this scenario, the state and energy MSE of bNN becomes comparable to that of HNN. Meanwhile, pHNN continues to capture and evolve the non-harmonic force accurately, demonstrating superior performance in initial state tests compared to other models.

The problem with the presented neural networks is that in the real world, such neurons are noisy, and the output is a probabilistic input function. Therefore, in the appendix, we analyze some stochastic representations of the NNs.

According to [14], we consider the structure of a port-Hamiltonian system of the form

$$\begin{cases} \dot{x} = [J(x) - R(x)] \frac{\partial H}{\partial x}(x) + G(x)u, \\ \dot{y} = G^\top(x) \frac{\partial H}{\partial x}(x), \end{cases} \quad (42)$$

with $x \in \mathbb{R}^n$ the energy variable, $H(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ the total stored energy, $u, y \in \mathbb{R}^m$ the port power variables.

In this context, we can explore Evolution Strategies (ES) as a stochastic optimization technique to evolve a population of solutions. ES includes basic concepts such as mutation, selection, and fitness evaluation. These strategies simulate collective learning processes within a population of individuals.

ES is part of a broader category of evolutionary algorithms, including genetic algorithms (GA) and evolutionary programming. These methods start with a random population and evolve through selection, mutation, and recombination processes. The interaction between creating new genetic information and its evaluation and selection drives evolution. Individuals who perform better have an increased chance of survival and produce offspring that inherit their genetic information.

In ES, the control synthesis problem is simplified to finding the minimum of a function, called fitness, within a feasible set of values. Each individual in the population is characterized by a set

of exogenous parameters π and endogenous parameters σ representing points in the search space and independent paths in the phase space, respectively. Sufficient initial conditions are provided to avoid overfitting and excessive computational cost.

The goal is for the set of trajectories to approach an equilibrium point under the influence of a stabilizing controller. The fitness metric evaluates the performance of the controller, which is derived from the energy function H_a . Mutation and self-adaptation act independently on each individual, generating new parameters accordingly:

$$\pi_{i+1} = \pi_i + \sigma_i \cdot N(0, 1)$$

$$\sigma_{i+1} = \sigma_i \cdot \exp(\tau \cdot N(0, 1)) + \tau' \cdot N(0, 1).$$

Here, τ and τ' are proportional to $(\sqrt{2n_\pi})^{-1}$ and $(\sqrt{2\sqrt{n_\pi}})^{-1}$, respectively, where n_π represents the dimension of the search space. These equations govern the evolution of the parameters within the ES framework and facilitate the exploration of the solution space.

6 Conclusions

The authors of [14] introduced a method that uses a neural network as an approximator to generate a gradient field for controller design, thereby avoiding the need to solve partial differential equations. To extend these results to a stochastic port-Hamiltonian scenario, it's necessary to introduce stochastic elements into the existing framework. This can be done by the following steps:

1. Define a stochastic version of the port-Hamiltonian system by introducing stochastic differential equations (SDEs) to model the system dynamics. For example, consider a system described by

$$dx = (J(x) - R(x)) \nabla H(x) dt + G(x)u dt + \Sigma(x) dW,$$

Where $\Sigma(x)$ is a matrix characterising the intensity of stochastic perturbations and dW is a standard Wiener process.

2. Extend the control synthesis methodology to account for the stochastic nature of the system. Modify the control law to ensure stability and performance despite stochastic disturbances. The control law might take the form

$$u = -K(x) + \mu(x)$$

where $K(x)$ is derived from the deterministic part of the system (as in the original paper) and $\mu(x)$ is an additional term to counteract stochastic disturbances.

Accordingly, an appropriate Lyapunov function must be identified or constructed to ensure the stability of the stochastic system. Lyapunov analysis for stochastic systems typically uses Itô's formula and aims to show that the expected value of the Lyapunov function decreases with time. Using evolutionary strategies, adapt the optimization process to handle the stochastic nature of the system. This may involve modifying the fitness function to account for stochastic behavior, possibly by considering expected performance over a range of stochastic disturbances. Finally, a new case study will extend the ball and beam system to include stochastic elements. Simulations can be used to demonstrate the stochastic disturbances.

A Stochastic Neural Networks

Neurons, the basic computational units in the brain, are known to form neural networks (NNs) by connecting through synaptic junctions. However, real neurons introduce noise that makes the output a probabilistic function of the input. Unlike deterministic NNs, the activation map is stochastic. These networks are better at avoiding local minima during training. They are better suited to tasks with noisy or incomplete data, although implementation and training complexity increases in the stochastic case. In particular, their randomness improves their robustness against overfitting since they do not learn the noise in the training data precisely as deterministic networks do.

The interconnection of neurons from different layers creates a neural network. The primary components for the synapses and neurons that make up a stochastic neural network are the stochastic neurons implemented by magnetic tunnel junctions (MTJs). Without going into the physical details, it can be described that an MTJ device consists of two ferromagnets separated by a thin insulator and is characterized by a switching (activation) probability (see Zhu and Park [9] for more technical details). In the artificial neural network model, input spikes flow into synapses, each assigned a weight, as shown in Figure 5.

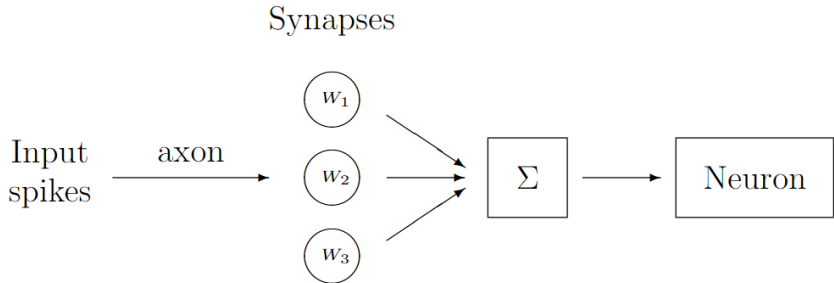


Figure 5: Schematic representation of a basic artificial Neuron block.

Then

$$\Sigma := \text{Input} \times \text{Weights} = \text{Summed Output}$$

defines the weighted summation and finally the neuron fires depending on a threshold or activation function. Vreeken [10] introduced *Spiking Neural Networks* (SNNs), artificial neural networks designed to mimic natural neural networks closely. Neurons emit these short bursts of electrical energy when they accumulate in sufficient numbers. The spike travels along the neuron’s axon and is controlled by the synapse, which consists of the end of the axon, a synaptic gap, and the initial part of the dendrite. Neurons must then use the spatial and temporal information of incoming spike patterns to encode their message to other neurons.

A modification of stochastic neural networks is presented by Yu et al. [11] with a *Simple and Effective Stochastic Neural Network* (SE-SNN). This approach is based on the notion that activation uncertainty can be modeled at each layer by predicting a Gaussian mean and variance and then sampling the layer during the forward pass. This method emphasizes the activation distribution and incorporates an activation regularizer to optimize models with high uncertainty while still predicting the target label. SE-SNN exhibits favorable properties in pruning, adversarial defense, learning with label noise, and improving model calibration. In the context described, instead of encoding data using binary numbers, we can consider the representation by the probability of encountering 1’s in bit streams. Stochastic Computing (SC) uses this technique to represent continuous values using streams of random bits. It uses conventional digital logic to perform computations based on stochastic bit

streams. SC, as described in [8], encompasses a wide range of techniques, but the focus here is on a specific aspect used in [7]. The main benefits of SC include reducing hardware complexity and enabling fault-tolerant computing. This reduction in complexity comes from implementing functions such as the sigmoid, hyperbolic tangent, and exponential functions using linear finite state machines, resulting in lower hardware costs. However, this may result in reduced computational accuracy. In addition, as noted in [7], the noise introduced in SC can mitigate overfitting problems and improve inference accuracy. Consequently, both input and output are represented by bit streams, with values encoded as probabilities of encountering 1's in these streams. This approach emphasizes the fusion of probabilistic methods with digital computation, providing a unique data representation and processing paradigm.

References

- [1] Van Der Schaft, Arjan, and Dimitri Jeltsema. "Port-Hamiltonian systems theory: An introductory overview." *Foundations and Trends in Systems and Control* 1.2-3 (2014): 173-378.
- [2] Cordoni, Francesco, Luca Di Persio, and Riccardo Muradore. "Stochastic port-Hamiltonian systems." *Journal of Nonlinear Science* 32.6 (2022): 91.
- [3] Desai, Shaan A., et al. "Port-Hamiltonian neural networks for learning explicit time-dependent dynamical systems." *Physical Review E* 104.3 (2021): 034312.
- [4] Talasila, Viswanath, J. Clemente-Gallardo, and A. J. Van der Schaft. "Discrete port-Hamiltonian systems." *Systems & Control Letters* 55.6 (2006): 478-486.
- [5] Kotyczka, Paul, and Laurent Lefevre. "Discrete-time port-Hamiltonian systems: A definition based on symplectic integration." *Systems & Control Letters* 133 (2019): 104530.
- [6] Talasila, Viswanath, J. Clemente-Gallardo, and A. J. van der Schaft. "Discrete port-Hamiltonian systems: mixed interconnections." *Proceedings of the 44th IEEE Conference on Decision and Control*. IEEE, 2005.
- [7] Liu, Yidong, et al. "A survey of stochastic computing neural networks for machine learning applications." *IEEE Transactions on Neural Networks and Learning Systems* 32.7 (2020): 2809-2824.
- [8] Brown, Bradley D., and Howard C. Card. "Stochastic neural computation. I. Computational elements." *IEEE Transactions on computers* 50.9 (2001): 891-905.
- [9] Zhu, Jian-Gang Jimmy, and Chando Park. "Magnetic tunnel junctions." *Materials today* 9.11 (2006): 36-45.
- [10] Vreeken, Jilles. "Spiking neural networks, an introduction." (2003).
- [11] Yu, Tianyuan, et al. "Simple and effective stochastic neural networks." *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. No. 4. 2021.
- [12] Ehrhardt, Matthias, Thomas Kruse, and Antoine Tordeux. "The Collective Dynamics of a Stochastic Port-Hamiltonian Self-Driven Agent Model in One Dimension." to appear in *ESAIM: Mathematical Modelling and Numerical Analysis*, 2024.

- [13] Gardiner, Crispin W. *Handbook of stochastic methods*. Vol. 3. Berlin: Springer, 1985.
- [14] Colonius, Fritz, and Lars Grüne, eds. *Dynamics, Bifurcations and Control*. Vol. 273. Springer Science & Business Media, 2002.
- [15] Wang, Chuwei, Shanda Li, Di He, Liwei Wang. "Is L^2 physics informed loss always suitable for training physics informed neural network?" *Adv. Neural Inform. Process. Syst.* 35 (2022): 8278-8290.