



OPEN ACCESS

EDITED BY
Emmanuel Benjamin Vander Poorten,
Faculty of Engineering Technology, KU
Leuven, Belgium

REVIEWED BY
Chaoyang Shi,
Tianjin University, China
Di Wu,
KU Leuven, Belgium

*CORRESPONDENCE
Genmeng Wang,
✉ genmeng.wang@insa-lyon.fr

RECEIVED 16 December 2025
REVISED 03 April 2026
ACCEPTED 08 April 2026
PUBLISHED 18 May 2026

CITATION

Wang G, Chalard R, Cifuentes J and
Pham MT (2026) Fast adaptation of
physics-informed hybrid models for
pneumatic artificial muscles.
Front. Robot. AI 13:1769141.
doi: 10.3389/frobt.2026.1769141

COPYRIGHT

© 2026 Wang, Chalard, Cifuentes and
Pham. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Fast adaptation of physics-informed hybrid models for pneumatic artificial muscles

Genmeng Wang^{1*}, Remi Chalard², Jenny Cifuentes³ and
Minh Tu Pham¹

¹University Lyon, INSA Lyon, Université Claude Bernard Lyon 1, Ecole Centrale de Lyon, CNRS, Ampère, UMR5005, Villeurbanne, France, ²Laboratory IBISC, Evry University, Paris-Saclay, Evry Cedex, France, ³ICADE, Department of Quantitative Methods, Faculty of Economics and Business Administration, and the Institute for Research in Technology (IIT), ICAI School of Engineering, Universidad Pontificia Comillas, Madrid, Spain

Recent advances in foundation models and physics-informed neural networks have demonstrated remarkable generalization and adaptation capabilities across diverse domains. Inspired by these properties, we investigate the adaptation potential of a previously proposed physics-informed hybrid model (PIHM) designed for pneumatic artificial muscles (PAMs). Through a series of experiments, it is demonstrated that, by incorporating an adapter based on physical prior knowledge, the PIHM model can be fine-tuned to transfer across different entity types while significantly reducing training time and maintaining competitive accuracy. The optimization efficiency of the proposed adapter has also been validated through comparison with other transfer learning techniques, such as full fine-tuning (FFT), partial fine-tuning (PFT), and low-rank adaptation (LoRA). These results suggest that embedding structured prior knowledge within hybrid architectures offers a promising solution for fast adaptation of PIHMs in dynamic system modeling.

KEYWORDS

fine-tuning, foundation model, model adaptation, physics-informed neural networks, pneumatic artificial muscles, transfer learning

1 Introduction

Since the Center for Research on Foundation Models (CRFMs) in the Human-Centered Artificial Intelligence (HAI) Institute of Stanford University coined the term “foundation model” in August 2021 (Bommasani et al., 2021), a growing body of research has been conducted on foundation models in various domains, including natural language processing and robotics manipulation. The well-known foundation models, such as BERT (Devlin et al., 2019) and GPT-4 (Achiam et al., 2023) in language processing, DALL-E (Ramesh et al., 2021) in image generation, and RT-X and $\pi_{0.5}$ vision–language–action (VLA) models in robotics, have two common features: they are trained on broad data and can be adapted to a wide range of tasks.

The aim of the foundation model approach is to develop a general model that captures some “knowledge” based on large-scale, sometimes multi-modal, information. The “knowledge” learned from the source tasks is expected to be further adapted to downstream tasks of interest without the need for training from scratch (Zhou et al., 2024; Brohan et al., 2023; Bommasani et al., 2021). The adaptation of foundation models can be achieved through different

methods, such as full fine-tuning (FFT), partial fine-tuning (PFT), and parameter-efficient fine-tuning (PEFT) (Houlsby et al., 2019; He et al., 2022; Lv et al., 2023; Wang L. et al., 2025).

A similar trend toward model reuse and efficient adaptation has emerged in the field of physics-informed neural networks (PINNs). PINNs provide a promising framework for solving partial differential equations (PDEs) by integrating neural networks with governing physical laws (Raissi et al., 2017). Alongside their development, numerous studies have explored strategies to adapt pre-trained PINNs to new PDE configurations, aiming to avoid costly re-optimization when system parameters or boundary conditions change (Desai et al., 2021; Prantikos et al., 2023; Wang Y. et al., 2025). In contrast to foundation models, which emphasize large-scale pre-training on massive datasets to achieve broad generality, adaptation in PINNs targets data-efficient specialization, exploiting embedded physical priors to rapidly adjust to closely related problems using limited new information (Torres and Niepert, 2024).

Based on evidence from both foundation models and PINNs that pre-learned representations can be effectively transferred across domains, platforms, and tasks, we examine the adaptation capability of a previously proposed physics-informed hybrid model (PIHM) for pneumatic artificial muscle (PAM) modeling. This hybrid model, proposed by the same authors in 2024 for modeling the quasi-static and thermodynamic behavior of PAM, has shown promising performance compared to traditional analytical models and pure data-driven models, such as feedforward neural networks and recurrent neural networks. More importantly, due to the presence of physical prior knowledge, this hybrid model has also demonstrated good generalization to unseen scenarios within the target system (Wang et al., 2025a; b). In this study, previous findings are extended by probing the adaptation capacity of the same hybrid model to new entity types. An adapter based on prior physics knowledge is introduced to augment the parameter space of a pre-learned PIHM for fine-tuning. The objective is to evaluate whether the underlying structure and physical guidance of the hybrid model facilitate fast adaptation to new dynamic systems, thus aligning with the broader goals of generalization and sample efficiency associated with foundation models.

The remainder of this article is organized as follows: Section 2 provides a review of related work on foundation models, hybrid approaches, and the modeling of PAMs. Section 3 thoroughly describes the proposed algorithm for fast adaptation. Section 4 presents the preparation of the dataset and the experimental setup. This section also presents the results obtained by applying the fast adaptation algorithm to two entity types. Finally, Section 5 summarizes the key findings and discusses potential future research directions.

2 Preliminaries/backgrounds/related work

2.1 Pneumatic artificial muscles

McKibben PAMs are bio-inspired actuators designed to mimic the function of human skeletal muscles. Due to their muscle-like behavior, along with higher power-to-weight and power-to-volume

ratios than those of electric and hydraulic actuators, PAMs are ideal for the development of rehabilitation devices and bio-robotics (Noritsugu and Tanaka, 1997; Park et al., 2014; Ding et al., 2008; Yaxi and Xu, 2021; Kalita et al., 2022).

There are several models in existence for the analysis or control of PAMs, including static and thermodynamic models. A quasi-static model of PAMs focuses on capturing the relationship between the actuator's internal pressure P , length variation ϵ , and the generated axial force F . The development of the quasi-static model usually starts by assuming that the PAM is a perfect cylinder and the thread length b remains constant during the contraction. Then, the equilibrium of virtual work \mathbf{W} done by the inner pressure P and external force F is denoted as shown in Equations 1–3:

$$d\mathbf{W}_{in} = d\mathbf{W}_{out}, \quad (1)$$

$$d\mathbf{W}_{in} = P dV, \quad d\mathbf{W}_{out} = -FdL, \quad (2)$$

$$P dV = -FdL, \quad (3)$$

where dL is the change in length and dV is the change in volume. Researchers have incorporated different estimations of dV/dL into Equation 3 to develop their respective quasi-static model (Chou and Hannaford, 1996; Tondu and Lopez, 2000; Bou Saba et al., 2016). Moreover, a low-order polynomial approximation of dV/dL was employed by Hildebrandt et al. (2003).

The development of a thermodynamic model typically involves the application of the first law of thermodynamics to the gas chamber, relating mass flows \dot{m}_{in} and \dot{m}_{out} . To streamline the modeling process, a number of simplifying assumptions are usually employed: (i) the gas is considered to behave as an ideal gas, (ii) the pressure and temperature are assumed to be uniformly distributed within the chamber, (iii) kinetic and potential energy contributions are deemed negligible, (iv) the process is considered adiabatic, and (v) the incoming flow is already at the temperature of the gas in the chamber. In view of these conditions, the time derivative of the internal pressure \dot{P} within the PAM is hereby defined as follows (Equation 4):

$$\dot{P} = \frac{kRT}{V} (\dot{m}_{in} - \dot{m}_{out}) - k \frac{\dot{V}}{V} P, \quad (4)$$

where $k = 1.4$ is the specific heat ratio of air, $R = 287\text{J}/(\text{kg} \cdot \text{K})$ is the gas constant for dry air, and $T = 293\text{K}$ is the temperature. This model can be rewritten by inverting the process and adding the mass flow $\dot{m} = \dot{m}_{in} - \dot{m}_{out}$ (Equation 5):

$$\dot{m} = \frac{\dot{P}V}{\gamma RT} + \frac{P\dot{V}}{RT}. \quad (5)$$

This simplified thermodynamic model, first introduced by Richer and Hurmuzlu (2000), has been widely adopted in research related to PAMs.

It is noteworthy that the key point of the modeling, in both the static and thermodynamic models, is the representation of the volume of the PAM. In various studies, the representation of the PAM volume is either based on a geometrical model (Chou and Hannaford, 1996; Tondu and Lopez, 2000; Bou Saba et al., 2016) or approximated using a polynomial (Itto and Kogiso, 2011; Hošovský et al., 2015; Hildebrandt et al., 2003).

Nevertheless, the strong nonlinear and hysteretic behavior exhibited by PAMs indicates that the accuracy of existing models can still be enhanced. The previously proposed PIHM approach learns to estimate the volume and the rate of change of volume of PAMs from data, with the guidance of an analytical model (Equation 3 or Equation 5), and has shown encouraging results in PAM modeling.

2.2 Physics-informed hybrid model

In this study, the term “hybrid” refers to the integration of an NN model with an analytical model. A number of hybridization methods have been proposed in the literature, and each method makes use of the strengths of both methodologies (Raissi et al., 2017; Lutter et al., 2019; Hošovský et al., 2015; Gaskin et al., 2023). Among the existing strategies, an approach tailored to the specific challenges of PAM modeling has been developed following the physics-informed paradigms to address their quasi-static and thermodynamic behaviors. This model had been previously proposed by the authors in earlier studies and is now revisited in light of its promising adaptation capabilities and performance across unseen conditions (Wang et al., 2025a; b).

This PIHM of PAMs is designed to estimate the gas chamber volume V from a set of physical input variables. Unlike traditional PINN methods, whose optimization relies on a loss function comprising both physical and data components, the PIHM approach requires only the physical component of the loss function, rendering it more feasible for practical applications. This hybrid model provides good approximations of V and \dot{V} with the guidance of an analytical model and also demonstrates good generalization to unseen trajectories of the same system. Specifically, the neural network receives $\mathbf{x} = [\varepsilon, \dot{\varepsilon}, P, \dot{P}]$ as input, and subsequently, an analytical model as shown in Equation 3 or Equation 5 handles the estimated volume \hat{V} , as well as its partial derivative with respect to the input contraction ratio ε and inner pressure P to form the final prediction $y = \hat{m}$ or $y = \hat{F}$. One can develop a reliable solution for modeling of PAMs following Algorithm 1. It is worth noting that steps 8–12 correspond to the Adam optimizer, a widely adopted stochastic gradient-based optimization algorithm designed for deep neural networks (Kingma and Ba, 2014).

2.3 Adaptation in PINNs

Adaptation in PINNs refers to the reuse of pre-trained models to efficiently solve new but related physical problems, such as PDEs with modified parameters and boundary conditions. Compared to retraining a PINN from scratch, the aim of the adaptation methods is to leverage previously learned solution structures and embedded physical priors to accelerate convergence and reduce computational costs.

Interest in this topic stems from the scalability limitations of traditional PINNs: despite their strong physical consistency, even minor modifications to the problem specification often require complete re-optimization. This constraint has driven the development of adaptation strategies to enhance data efficiency,

Require: Dataset $\mathcal{D} = \{\mathbf{x}^{(n)}, y^{(n)}\}_{n=1}^N$, hyperparameters, *Thermodyna*

Initialize neural network parameters θ

for all epochs **do**

for all $(\mathbf{x}_k, \mathbf{y}_k), k \leftarrow 1:K$ **do**

 ▷ k^{th} batch from \mathcal{D}

$\hat{\mathbf{V}}_k \leftarrow \phi(\mathbf{x}_k; \theta)$

 ▷ Apply forward propagation

if *Thermodyna* **then**

 ▷ Thermodynamic model

$\hat{\mathbf{V}}_k \leftarrow \frac{\partial \hat{\mathbf{V}}_k}{\partial \varepsilon} \hat{\varepsilon}_k + \frac{\partial \hat{\mathbf{V}}_k}{\partial P} \hat{P}_k$ ▷ Compute derivative of $\hat{\mathbf{V}}_k$ using automatic differentiation

$\hat{\mathbf{y}}_k \leftarrow \frac{P_k \hat{\mathbf{V}}_k}{\gamma RT} + \frac{P_k \dot{\hat{\mathbf{V}}}_k}{RT}$

 ▷ Predict mass flow with analytical model

else

 ▷ Quasi-Static Model

$\frac{d\hat{\mathbf{V}}_k}{dL} \leftarrow \frac{\partial \hat{\mathbf{V}}_k}{\partial \varepsilon L_0}$ ▷ Compute derivative of $\hat{\mathbf{V}}_k$ using automatic differentiation

$\hat{\mathbf{y}}_k \leftarrow P_k \frac{d\hat{\mathbf{V}}_k}{dL}$

 ▷ Predict force with the analytical model

end if

$\mathcal{L}_k \leftarrow \text{MSE}(\mathbf{y}_k, \hat{\mathbf{y}}_k) + \frac{\lambda}{2} \|\theta\|_2^2$

 ▷ Compute loss

$\mathbf{g}_k \leftarrow \nabla_{\theta} \mathcal{L}_k + \lambda \theta_{k-1}$

 ▷ Compute gradient

$M_k \leftarrow \beta_1 M_{k-1} + (1 - \beta_1) \mathbf{g}_k$

 ▷ Update biased first moment

$V_k \leftarrow \beta_2 V_{k-1} + (1 - \beta_2) \mathbf{g}_k^2$

 ▷ Update biased second moment

$\hat{M}_k \leftarrow \frac{M_k}{1 - \beta_1^k}, \hat{V}_k \leftarrow \frac{V_k}{1 - \beta_2^k}$

 ▷ Correct bias

$\theta_k \leftarrow \theta_{k-1} - \eta \frac{\hat{M}_k}{\sqrt{\hat{V}_k + \epsilon}}$

 ▷ Update parameters

end for

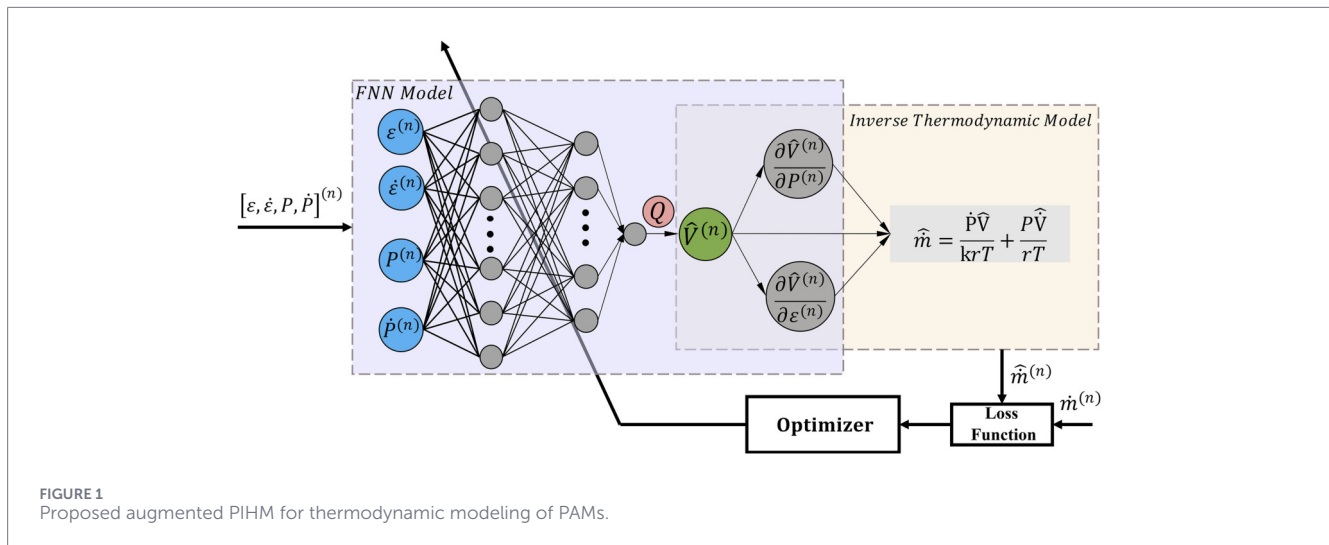
end for

return trained parameters θ^*

Algorithm 1. Training procedure of physics-informed hybrid model of PAMs.

thereby enabling PINNs to be redeployed to new PDEs with fewer data and computational demands.

To achieve adaptation and transfer learning in PINNs, several techniques have been proposed. Full fine-tuning involves using a pre-trained PINN from a source problem as the base model, followed by optimizing all network parameters for the target problem. Although this approach is effective for closely related tasks, it remains computationally intensive (Penwarden et al., 2023). Partial or lightweight fine-tuning freezes a subset of pre-trained parameters, updating only selected layers to reduce training costs while preserving transferable physical representations (Chakraborty, 2021). Recently, parameter-efficient adaptation methods such as low-rank adaptation (LoRA) have been introduced to the PINN domain (Hu et al., 2021; Majumdar et al., 2023). These techniques inject a small number of trainable parameters



into the frozen pre-trained network, enabling rapid and stable adaptation while preserving the original model architecture.

3 Fast adaptation of physics-informed hybrid models of PAMs

A key property of PIHM is that its learning process is guided by an analytical model; hence, the learned dynamics are derived from the physics prior. In this study, we focus on transferring the learned dynamics to unseen PAM by adding an adapter, which is also based on physics priors and therefore reduces the time and data needed when modeling new PAMs.

In this section, an adapter Q is introduced as a physical principle-based tuning mechanism for fast adaptation of a pre-trained PIHM. Given a pre-trained PIHM on a target PAM Σ_0 , we can obtain Equation 6:

$$y = f_{\Sigma_0}(\mathbf{x}; \theta^*), \quad (6)$$

where $\mathbf{x} = [\varepsilon, \dot{\varepsilon}, P, \dot{P}]$, $y = \hat{m}$ for thermodynamic modeling, $y = \hat{F}$ for quasi-static modeling, and $\hat{\cdot}$ refers to estimation. The pre-trained PIHM consists of two parts: the NN part for predicting V , and the analytical part ensuring that learned parameters are consistent with the laws of physics. This physical adapter Q is designed to encode additional physical prior knowledge into pre-trained models, enabling a fast adaptation to new systems. This proposed method augments the learned PIHM parameter solely by one dimension, where physical priors can be stored as empirical guidance (Equation 7):

$$\tilde{\Theta} = \{(\theta, Q) \mid \theta \in \Theta, \Theta \subset \mathbb{R}^d, Q \in \mathbb{R}\}. \quad (7)$$

In scenarios where fast adaptation of f_{Σ_0} to the new target PAM system Σ_1 is required and the dataset from Σ_1 is small, the original parameter space of the network parameters Θ is proposed to be extended to $\tilde{\Theta}$ by adding an adapter Q , which may also be considered a scalar multiplier. As illustrated in Figure 1, it is proposed to modify the pre-trained PIHM by adding an adapter

Q to its output layer. The initial guess of Q can be chosen by roughly dividing the initial volume V_1 of Σ_1 by the initial volume V_0 of Σ_0 using the formula for the volume of a cylindrical gas chamber $V = \frac{\pi D^2 L}{4}$.

Algorithm 2 outlines the step-by-step implementation of the proposed adaptation method. By providing a dataset \mathcal{D}_{FA} , which is much smaller than \mathcal{D} required in Algorithm 1, one can follow the steps in Algorithm 2 to fine-tune a pre-trained PIHM to adapt to a different PAM Σ_1 . The tuning process is performed on all pre-trained parameters, and the initially guessed Q with different learning rates as clarified in line 19, Algorithm 2. The learning rate for Q is scaled by a factor of 0.2 compared to the reused neural network parameters, aiming to focus fine-tuning efforts on pre-trained parameters rather than adapters—as the primary behavioral patterns are assumed to be embodied by the pre-trained parameters.

4 Experimental results

This section presents the evaluation of the proposed fast adaptation method in thermodynamic modeling of two novel PAM entities, as shown in Figures 2B,C. A PIHM model ϕ_A is pre-trained with data collected from Figure 2A; then, adapters are added to ϕ_A for model adaptation, as illustrated in Algorithm 2. The obtained models are noted as ϕ_B^{FA} and ϕ_C^{FA} , where FA stands for fast adaptation. Other transfer learning strategies such as FFT, PFT, and LoRA were also applied in Figures 2B,C to better demonstrate the positioning of the proposed adaptation approach. These adapted models are noted as ϕ_B^{FFT} , ϕ_B^{PFT} , ϕ_B^{LoRA} , ϕ_C^{FFT} , ϕ_C^{PFT} , and ϕ_C^{LoRA} , where the superscripts denote the methods and the subscripts denote the target objects.

The test bench platform used in the experimental procedure is shown in Figure 3. This platform is mounted on the experimental stand and incorporates a pair of antagonistically connected PAM actuators to enable rotational motion of the right-hand component about its axis. There is also a stable high-pressure gas reservoir capable of supplying 6 bars to both PAMs, with this gas supply

Require: Dataset $\mathcal{D}_{FA} = \{\mathbf{x}^{(n)}, y^{(n)}\}_{n=1}^{N_{FA} \ll N}$, trained parameters θ^* , hyperparameters, *Thermodyna*

Initial guess of Q

- ▷ Based on physical prior

$$\hat{\theta} \leftarrow (\theta, Q)$$

- ▷ Augmented parameters

$$\tilde{\phi}(\mathbf{x}; \hat{\theta}) \leftarrow Q \cdot \phi(\mathbf{x}; \theta^*)$$

- ▷ Transfer the pre-learned knowledge

for all epochs do

- for all $(\mathbf{x}_k, \mathbf{y}_k), k \leftarrow 1:K$ do**
- ▷ k^{th} batch from \mathcal{D}_{FA}
$$\hat{\mathbf{V}}_k \leftarrow \tilde{\phi}(\mathbf{x}_k; \hat{\theta})$$
- ▷ Apply forward propagation
- if *Thermodyna* then**
- ▷ Thermodynamic model
$$\hat{\mathbf{V}}_k \leftarrow \frac{\partial \hat{\mathbf{V}}_k}{\partial \boldsymbol{\varepsilon}} \boldsymbol{\varepsilon}_k + \frac{\partial \hat{\mathbf{V}}_k}{\partial P_k} P_k$$
- ▷ Compute derivative of $\hat{\mathbf{V}}_k$ using automatic differentiation
$$\hat{\mathbf{y}}_k \leftarrow \frac{P_k \hat{\mathbf{V}}_k}{\gamma RT} + \frac{P_k \hat{\mathbf{V}}_k}{RT}$$
- ▷ Predict mass flow with analytical model
- else**
- ▷ Quasi-static model
$$\frac{d\hat{\mathbf{V}}_k}{dL} \leftarrow \frac{\partial \hat{\mathbf{V}}_k}{\partial \boldsymbol{\varepsilon} L_0}$$
- ▷ Compute derivative of $\hat{\mathbf{V}}_k$ using automatic differentiation
$$\hat{\mathbf{y}}_k \leftarrow P_k \frac{d\hat{\mathbf{V}}_k}{dL}$$
- ▷ Predict force with the analytical model
- end if**
$$\mathcal{L}_k \leftarrow \text{MSE}(\mathbf{y}_k, \hat{\mathbf{y}}_k) + \frac{\lambda}{2} \|\hat{\theta}\|_2^2$$
- ▷ Compute loss
$$g_k \leftarrow \nabla_{\hat{\theta}} \mathcal{L}_k + \lambda \hat{\theta}_{k-1}$$
- ▷ Compute gradient
$$M_k \leftarrow \beta_1 M_{k-1} + (1 - \beta_1) g_k$$
- ▷ Update biased first moment
$$v_k \leftarrow \beta_2 v_{k-1} + (1 - \beta_2) g_k^2$$
- ▷ Update biased second moment
$$\hat{M}_k \leftarrow \frac{M_k}{1 - \beta_1^k}, \quad \hat{v}_k \leftarrow \frac{v_k}{1 - \beta_2^k}$$
- ▷ Correct bias
$$\theta_k \leftarrow \theta_{k-1} - \eta \frac{\hat{M}_k}{\sqrt{\hat{v}_k + \epsilon}}, \quad Q_k \leftarrow Q_{k-1} - 0.2 \cdot \eta \frac{\hat{M}_k}{\sqrt{\hat{v}_k + \epsilon}}$$
- ▷ Update parameters
- end for**
- end for**
- return** trained parameters $\hat{\theta}^*$

Algorithm 2. Fast adaptation of physics-informed hybrid model of PAMs.

controlled using two independent servo valves. The directly measurable values are rotation angle ρ , internal gas pressure P , and the servo valve control voltage u . Furthermore, the mass flow rate \dot{m} can be calculated using the characteristic mapping of the servo valves verified by Olaby et al. (2005). Similarly, the contraction ratio ε can also be determined through a kinematic model of the experimental platform.

4.1 Data preparation

Three training datasets \mathcal{D}_A , \mathcal{D}_B , and \mathcal{D}_C were prepared using a simple proportional derivative controller, each containing 20 trajectories sampled at a frequency of 1,000 Hz. These 20 trajectories were all obtained by tracing sinusoidal signals composed of five different frequencies [1, 2, 4, 10] *rad/s* and five different amplitudes [5, 10, 15, 20, 25] *degrees* in rotational axis of the test bench. By switching between different experimental PAM objects, \mathcal{D}_A , \mathcal{D}_B , and \mathcal{D}_C are acquired. By sparsifying the datasets \mathcal{D}_B and \mathcal{D}_C , one obtains the datasets \mathcal{D}_B^{FA} and \mathcal{D}_C^{FA} for model adaptation. Since the modeling objective of the experimental section is the thermodynamic behavior of the PAM, $\mathbf{x} = [\varepsilon, \dot{\varepsilon}, P, \dot{P}]$ serves as the model input, while $y = \dot{m}$ represents the model output. Although both PAMs were actuated during the experiments, only data from the upper muscle were used for model training as the two muscles were identical and demonstrated symmetrical dynamic behavior.

To further investigate whether the adapted PIHM model retains consistent generalization capabilities in unseen scenarios, two validation datasets were prepared for each target system. The first validation dataset is denoted as \mathcal{D}^{tri} , comprising 25 triangular trajectories composed of five frequencies [0.5, 1, 2, 4, 10] *rad/s* and five amplitudes [5, 10, 15, 20, 25] *degrees*. The second validation dataset is denoted as \mathcal{D}^{step} , comprising 10 step trajectories composed of 10 amplitudes [5, 10, 15, 20, 25, -5, -10, -15, -20, -25] *degrees*.

4.2 Training and fast adaptation

The workflow of fast adaptation is summarized in Figure 4. A general model was first trained using Algorithm 1 on \mathcal{D}_A , which is denoted as $\phi_A(\mathbf{x}; \theta_A^*)$. Concurrently, two baseline models were trained from scratch through Algorithm 1 using the \mathcal{D}_B and \mathcal{D}_C datasets, denoted as $\phi_B(\mathbf{x}; \theta_B^*)$ and $\phi_C(\mathbf{x}; \theta_C^*)$, respectively. The prediction accuracy and training time of these two baseline models are compared with the model obtained using the fast adaptation method, which are $\phi_B^{FA}(\mathbf{x}; \hat{\theta}_B^*)$ and $\phi_C^{FA}(\mathbf{x}; \hat{\theta}_C^*)$.

Before executing the fast adaptation of Algorithm 2, initial guesses for Q_B and Q_C must be drawn from physical knowledge to augment the pre-trained parameters θ_A^* into $\hat{\theta}_B$ and $\hat{\theta}_C$. As shown in Figure 2, the initial diameter of the air chamber in PAM in Figure 2B is twice that of PAM in Figure 2A, while their initial lengths are identical. The initial diameter of the air chamber in PAM in Figure 2C is twice that of PAM in Figure 2A, and its initial length is 2.25 times that of PAM in Figure 2A. Using the formula for the volume of a cylindrical gas chamber $V = \frac{\pi D^2 L}{4}$ and considering their initial diameter and length when not in service, one can choose initial guess $Q_B = 4$ and $Q_C = 9$. Based on the initial guesses provided, Algorithm 2 is capable of fine-tuning both the learned parameters θ_A^* and the augmented parameters Q to align them with the new PAM system. Furthermore, it is assumed that since the primary behavioral patterns have already been learned through the larger dataset \mathcal{D}_A , the dataset required for fine-tuning can be sparsified, thereby reducing the time and computational resources needed for knowledge transfer.

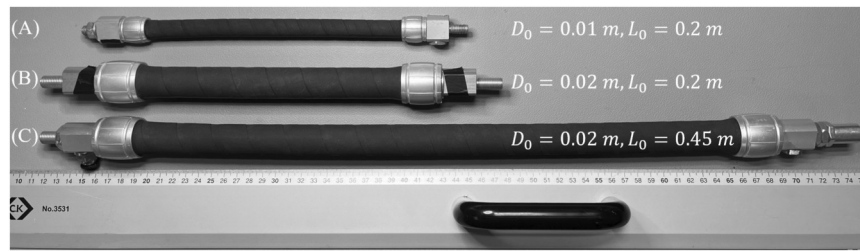
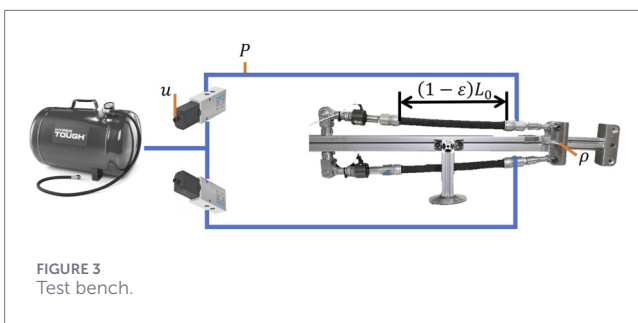


FIGURE 2

PAM objects in this study: (A) PAM with an initial length of 0.2 m and an initial diameter of 0.01 m; (B) PAM with an initial length of 0.2 m and an initial diameter of 0.02 m; (C) PAM with an initial length of 0.45 m and an initial diameter of 0.02 m.

FIGURE 3
Test bench.

The models ϕ_B^{FFT} and ϕ_C^{FFT} were obtained by fine-tuning all pre-trained parameters in ϕ_A on downstream datasets \mathcal{D}_B and \mathcal{D}_C . On the contrary, only the parameters in the first hidden layer are fine-tuned, while those in the subsequent two layers are frozen when applying the PFT strategy for ϕ_B^{PFT} and ϕ_C^{PFT} .

The LoRA method injects rank decomposition matrices into layers of a neural network while freezing the pre-trained model weights to reduce the amount of trainable parameters for downstream tasks. In this method, the forward calculation is modified from $h = W_0x$ to $h = W_0x + BAx$, where W_0 is a pre-trained matrix $W_0 \in \mathbb{R}^{d \times k}$, B is a matrix $B \in \mathbb{R}^{d \times r}$ initialized to 0, and A is a randomly initialized matrix $A \in \mathbb{R}^{r \times k}$. By ensuring $r \ll \min(d, k)$, one can save both computational time and resources for model adaptation to downstream tasks. In this study, two decomposition matrices $A_1 \in \mathbb{R}^{2 \times 128}$ and $B_1 \in \mathbb{R}^{4 \times 2}$ were applied to the first hidden layer of $\phi_A(x; \theta_A^*)$, while $A_2 \in \mathbb{R}^{8 \times 128}$ and $B_2 \in \mathbb{R}^{128 \times 8}$ were applied to the second hidden layer. The hyperparameters setup for all aforementioned models can be found in Table 1.

4.3 Results

A summary of performance metrics can be found in Table 2. The result shows that, with a sparsified dataset and pre-trained parameters, the model adaptation phase requires less time than training from scratch. It is also observed that the models obtained via the proposed fast adaptation approach achieve better prediction precision than the baseline models in the target task, which indicates

strong performance in downstream task adaptation. However, on the validation dataset, the model based on the proposed method ϕ^{FA} still maintains competitive performance compared to the baseline models ϕ , ϕ^{FFT} , ϕ^{PFT} , and ϕ^{LoRA} , indicating that the proposed adaptation method possesses acceptable generalization capabilities.

Under identical computational resources (CPU: Intel i7-1370P, GPU: NVIDIA RTX A500 Laptop), the average time needed for execution of Algorithm 1 is 195,294 s in this study. Meanwhile, the average time required for Algorithm 2 is 467 s, which indicates a significant increase in efficiency. To better appreciate this improvement, it is worth recalling that Algorithm 1 corresponds to full model training from scratch, optimizing the entire parameter set θ , whereas Algorithm 2 performs fast adaptation by optimizing the augmented parameters $\tilde{\theta}$, starting from the pre-trained model $\phi_A(x; \theta_A^*)$. It has been noticed that the LoRA method takes slightly more time during the adaptation phase; this may be due to the need for more computations during the forward prediction process.

Beyond the enormous savings in computational time, the fast adaptation model also realizes a smaller loss at the end of training. As shown in Figures 5A,B, the loss curve of the proposed method (in green) consistently remains below that of baseline training in the small inset diagram in the right corner. It can also be observed on the left side of Figures 5A,B that the proposed method begins the optimization with a lower initial loss. Since the FFT, PFT, and LoRA methods reuse pre-trained parameters during the adaptation process without any biases in the first prediction, their initial loss during the adaptation phase effectively depends on the substantial knowledge accumulated during the pre-training phase. It can, therefore, be concluded that the proposed adapter helps reduce the initial loss during the adaptation process through additional knowledge based on the initial guess of Q_B and Q_C . Given that our method converges to the minimum earlier than other adaptation methods, the proposed adaptation strategy has also demonstrated its optimization efficiency. Additionally, it is worth noting that our proposed method does not encounter an initial plateau phase, unlike the training from scratch did at the beginning in Figures 5A,B.

On the validation dataset, predictions made by the derived models indicate that despite a relatively long plateau period

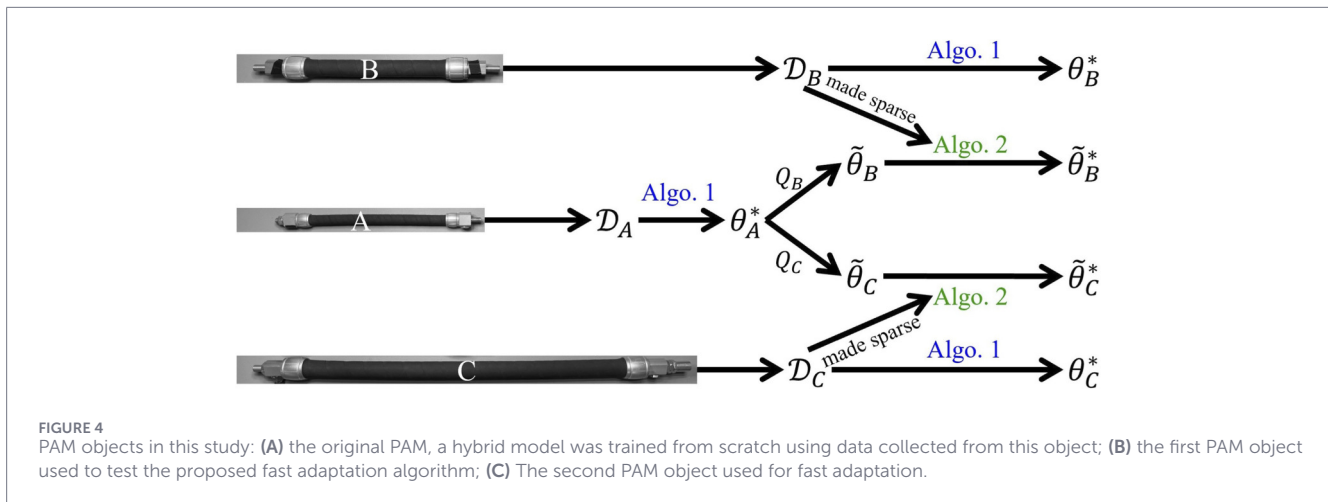


TABLE 1 Hyperparameters applied in this study.

Hidden layer	2×128
Activation function	SoftPlus
Learning rate	$1e^{-4}$
Weight decay	$1e^{-5}$
Optimizer	Adam
Loss function	MSE
Number of epochs	10,000
Batch size (training from scratch)	2,048
Batch size (adaptation)	128

following rapid descent, none of the four models ϕ_B , ϕ_C , ϕ_B^{FA} , and ϕ_C^{FA} encountered remarkable overfitting issues. The predictions by the fast adapted models and train from scratch models can be found in Figures 6, 7. The upper subplots of these two figures display results in the validation datasets \mathcal{D}_B^{tri} and \mathcal{D}_C^{tri} . Correspondingly, the lower subplots present results in the validation datasets \mathcal{D}_B^{step} and \mathcal{D}_C^{step} .

In Figure 6A, it can be observed that ϕ_B and ϕ_B^{FA} deliver nearly identical predictions when sharing the same inputs. Consequently, in Figure 7A, the divergence between the train from the scratch model and the adapted model becomes more pronounced, particularly evident during the negative half-cycle of second-high frequency trajectories. The L2 norm calculated from the validation dataset \mathcal{D}_B^{tri} in Table 2 supports these findings. The deviation between ϕ_B and ϕ_B^{FA} is smaller than the deviation between ϕ_C and ϕ_C^{FA} . A possible explanation may be attributed to geometric differences. The increased length of the PAM may result in a smaller-than-expected outflow as the distance between the air at the end of the gas chamber and outlet in the PAM in Figure 6C is longer than that in the PAM in Figure 6A. At relatively lower dynamic frequencies, this deviation is less pronounced because the required predicted gas flow rate is smaller. Similarly, at relatively higher motion frequencies, this phenomenon is not evident, possibly because the tension in the

antagonistically connected PAM compensates for it, hence, the observation is only noted in the trajectory at the second-highest frequency.

The train from scratch models and fast adapted models present similar biases in Figures 6B, 7B, with more noticeable deviations at peak values. The L2 norm calculated in Table 2 indicates that the models perform less favorably on the \mathcal{D}^{step} dataset than on the \mathcal{D}^{tri} dataset. A plausible reason is that the step trajectories in the validation dataset \mathcal{D}^{step} have fewer common features with the training dataset, thereby increasing the difficulty of model prediction. When confronted with step trajectories, ϕ_B and ϕ_B^{FA} perform slightly less favorably than ϕ_C and ϕ_C^{FA} , as shown in Table 2. This is presumed to the broader prediction range of \mathcal{D}_B^{step} than that of \mathcal{D}_C^{step} . This difference is illustrated in the Y-axis values between Figures 6B, 7B. A broader prediction range may encompass patterns not covered by \mathcal{D}_A , thereby increasing the difficulty of the prediction task depicted in Figure 6B.

Despite the performance divergence between ϕ and ϕ^{FA} on the validation datasets, this does not imply that the proposed fast adaptation approach is unsatisfactory. On the contrary, it demonstrates that this approach performs better when fine-tuning models to fit specific tasks, such as trajectory data used for fine-tuning.

The fast adapted model was subsequently deployed on a dSPACE MicroLabBox II platform for real-time control. With four times of forward calculations required, the control loop maintained a sampling rate of 1,000 Hz, which is suitable for our application. Moreover, as shown in Table 2, the number of parameters in the fast adapted model is nearly identical to that of the model trained from scratch, indicating nearly equivalent computational complexity. This highlights the advantage of the proposed adaptation method: it neither increases computational complexity nor expands the parameter space beyond a single dimension.

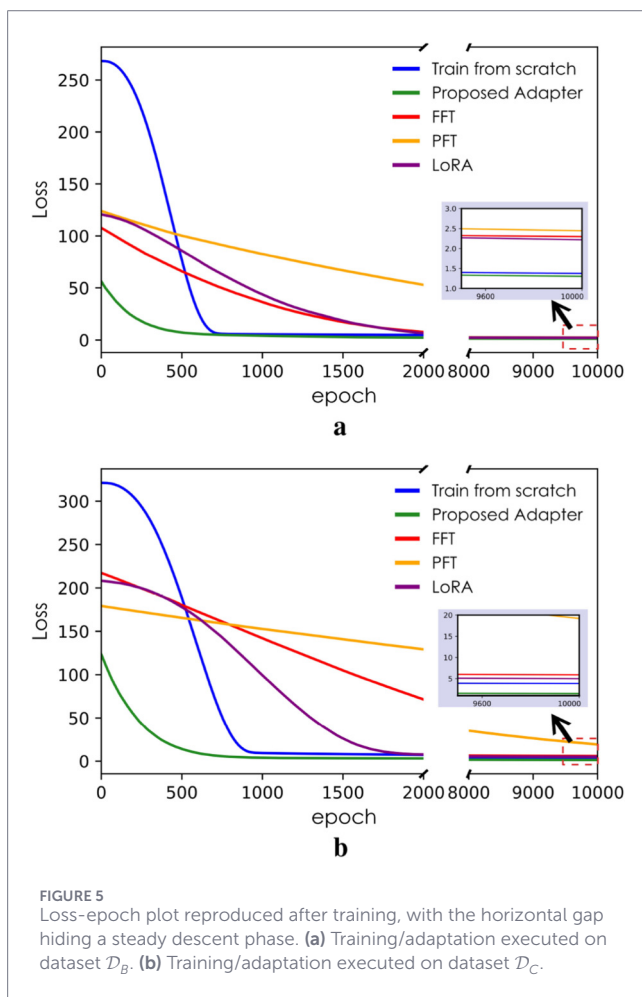
5 Conclusion

Based on two key characteristics of foundational models, the authors were motivated to investigate the adaptation capability

TABLE 2 Summary of performance metrics.

Target system	(B)					(C)				
Models	Train from scratch	FFT	PFT	LoRA	Ours	Train from scratch	FFT	PFT	LoRA	Ours
# Optimized parameters	17,281	17,281	640	1,288	17,282	17,281	17,281	640	1,288	17,282
L2-norm in target task	1.17	1.86	1.83	1.40	1.14	1.96	2.50	4.68	2.29	1.65
L2-norm in \mathcal{D}^{tri}	1.14	1.56	1.83	1.43	1.18	1.44	1.50	3.59	1.44	2.21
L2-norm in \mathcal{D}^{step}	3.42	3.38	3.43	3.61	3.60	2.47	2.55	3.46	2.50	2.61
Time used (s)	195,804	469	472	501	471	194,437	470	490	479	463
# training samples	31,464	315				30,952	310			

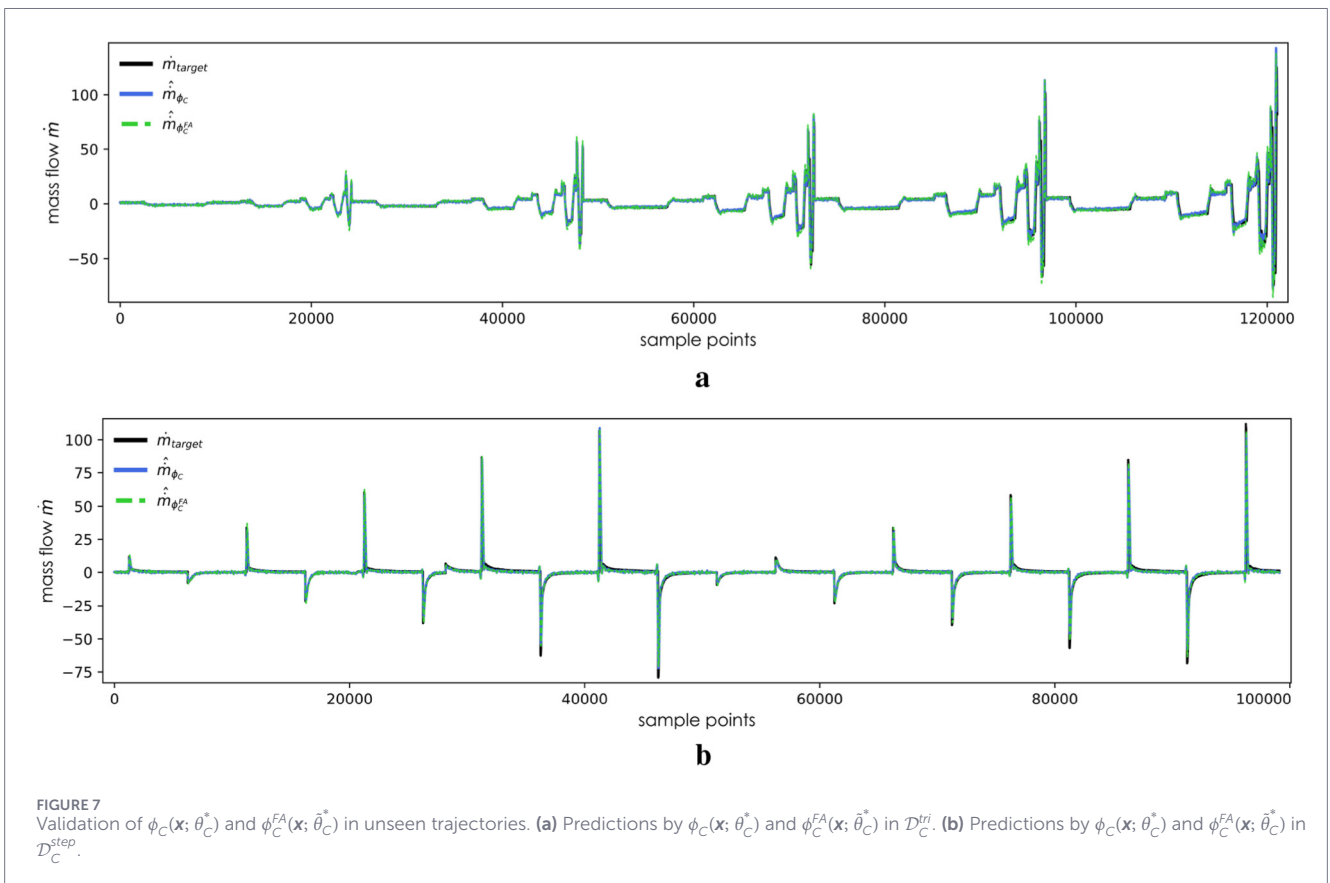
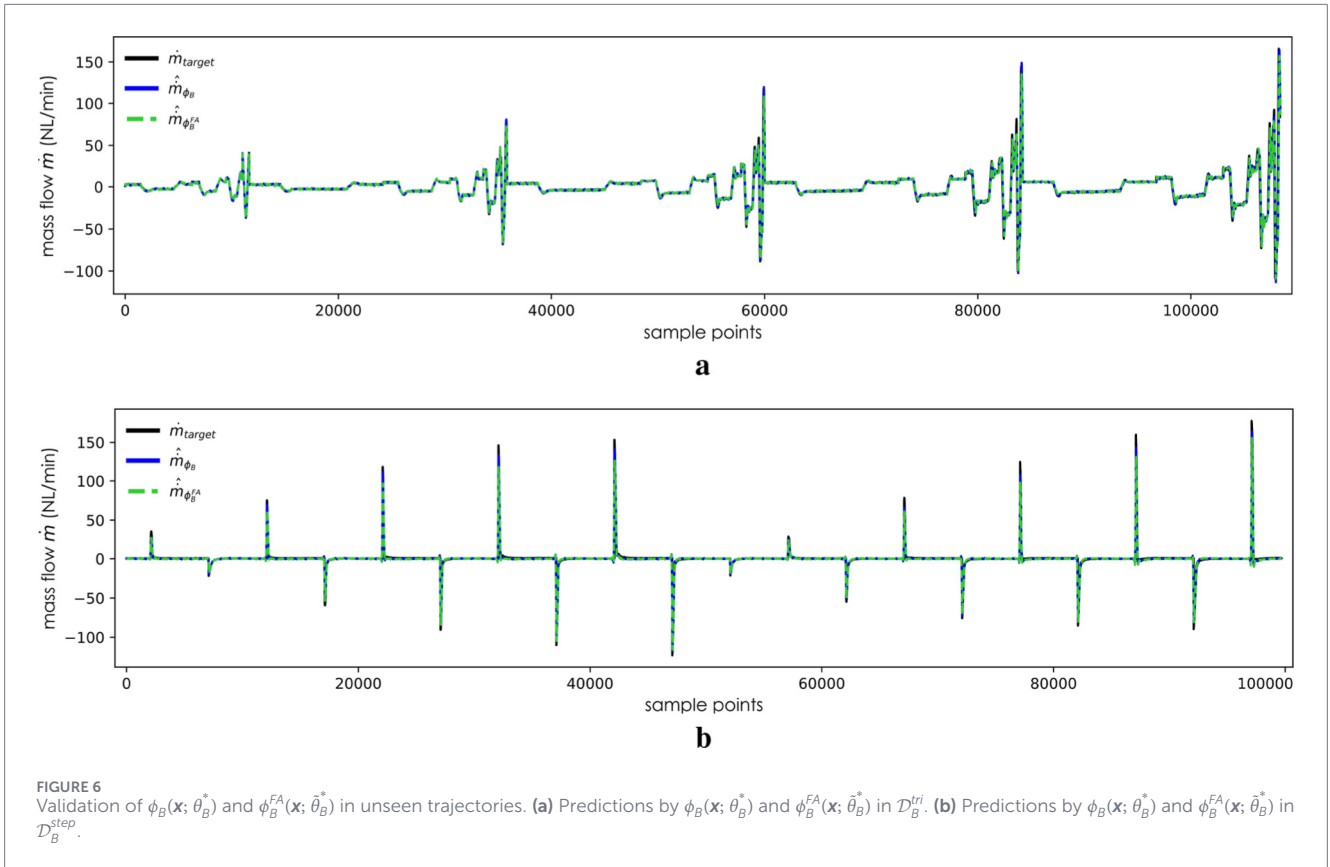
Values in bold indicate the best results obtained in each task after applying different adaptation methods to each target PAM object.



of their previously suggested physics-informed hybrid model for PAMs. In this regard, a study was conducted by comparing the baseline model with the fast adaptation model in terms

of computational time, estimation accuracy, and generalization capability to unseen scenarios. Three types of PAM entities were implemented in the experiments: one for pre-training and two for fine-tuning. The results demonstrate that incorporating human-knowledge-based adapters can significantly reduce computational and data requirements. Furthermore, the fast adaptation model outperformed the baseline models on the fine-tuning dataset but failed to fully match the performance of the baseline model on the validation dataset. This indicates that the model has acquired stronger task specificity, consistent with findings from other studies on foundation model fine-tuning. It is important to emphasize that this work is not intended to develop a foundational model since it was not trained on large-scale data and its generalization capabilities were evaluated on only a limited number of tasks.

These findings are regarded as offering a novel perspective for constructing foundational models for physics-informed representation. In PAM modeling, an intermediate physical quantity—namely, actuator volume—is introduced as a latent representation, while downstream applications may employ distinct analytical models to recover task-specific outputs such as force or airflow. A corresponding modeling approach exists within robotic systems: intermediate representations such as mass matrices serve as shared latent structures for both direct and inverse dynamics formulations. The selected intermediate variable has some relation when it comes across entities. For PAM models, the proposed adapter captures geometric scaling variations between actuators. Similarly, in robotic arm systems, geometric and inertial scaling effects arise due to variations in link lengths, masses, and degrees of freedom. Based on this structural correspondence, the aim of our ongoing study is to develop a more fundamental paradigm for physics-informed robotic arm models. These models utilize training datasets significantly larger than those employed in PAM studies, and the evaluation of the proposed adaptation paradigm can then be conducted across a broader spectrum of robotic scenarios.



Data availability statement

Publicly available datasets were analyzed in this study. These data can be found at <https://github.com/GenmengWANG/FastAdaptationOfPIHTM>.

Author contributions

GW: Data curation, Software, Conceptualization, Methodology, Investigation, Visualization, Validation, Formal analysis, Writing – review and editing, Writing – original draft. RC: Project administration, Funding acquisition, Methodology, Writing – review and editing. JC: Methodology, Writing – review and editing. MP: Methodology, Writing – review and editing, Supervision, Project administration, Funding acquisition.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships

References

- Achiam, O. J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., et al. (2023). *Gpt-4 technical report*.
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., et al. (2021). On the opportunities and risks of foundation models. *ArXiv*. doi:10.48550/arXiv.2108.07258
- Bou Saba, D., Bideaux, E., Brun, X., and Massioni, P. (2016). "A complete model of a two degree of freedom platform actuated by three pneumatic muscles elaborated for control synthesis," in BATH/ASME 2016 Symposium on Fluid Power and Motion Control. vol. V001T01A004 of ASME Proceedings. doi:10.1115/FPMP2016-1706V001T01A004
- Brohan, A., Brown, N., Carbajal, J., Chebotar, Y., Chen, X., Choromanski, K., et al. (2023). *RT-2: vision-language-action models transfer web knowledge to robotic control*. doi:10.48550/arXiv.2307.15818
- Chakraborty, S. (2021). Transfer learning based multi-fidelity physics informed deep neural network. *J. Comput. Phys.* 426, 109942. doi:10.1016/j.jcp.2020.109942
- Chou, C.-P., and Hannaford, B. (1996). Measurement and modeling of McKibben pneumatic artificial muscles. *IEEE Trans. Robotics Automation* 12, 90–102. doi:10.1109/70.481753
- Desai, S., Mattheakis, M., Joy, H., Protopapas, P., and Roberts, S. J. (2021). One-shot transfer learning of physics-informed neural networks. *ArXiv abs/2110.11286*. doi:10.48550/arXiv.2110.11286
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). "Bert: pre-training of deep bidirectional transformers for language understanding," in *North American chapter of the association for computational linguistics*.
- Ding, M., Ueda, J., and Ogasawara, T. (2008). Pinpointed muscle force control using a power-assisting device: system configuration and experiment. 181–186. doi:10.1109/BIOROB.2008.4762829
- Gaskin, T., Pavliotis, G. A., and Girolami, M. (2023). Neural parameter calibration for large-scale multiagent models. *Proc. Natl. Acad. Sci.* 120, e2216415120. doi:10.1073/pnas.2216415120
- He, K., Chen, X., Xie, S., Li, Y., Dollár, P., and Girshick, R. (2022). "Masked autoencoders are scalable vision learners," in 2022 IEEE/CVF Conference on Computer

that could be construed as a potential conflict of interest.

The author MP declared that they were an editorial board member of Frontiers at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Vision and Pattern Recognition (CVPR), 15979–15988. doi:10.1109/CVPR52688.2022.01553

Hildebrandt, A., Sawodny, O., Neumann, R., and Hartmann, A. (2003). "A cascaded tracking control concept for pneumatic muscle actuators," in 2003 European Control Conference (ECC), 2517–2522. doi:10.23919/ECC.2003.7085344

Hošovský, A., Pítel, J., and Židek, K. (2015). Enhanced dynamic model of pneumatic muscle actuator with Elman neural network. *Abstr. Appl. Anal.* 2015, 1–16. doi:10.1155/2015/906126

Houlsby, N., Giurgiu, A., Jastrzebski, S., Morrone, B., Laroussilhe, Q. D., Gesmundo, A., et al. (2019). "Parameter-efficient transfer learning for NLP," in Proceedings of the 36th International Conference on Machine Learning (PMLR), 2790–2799.

Hu, J. E., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., et al. (2021). Lora: low-rank adaptation of large language models. *ArXiv abs/2106.09685*. doi:10.48550/arXiv.2106.09685

Itto, T., and Kogiso, K. (2011). "Hybrid modeling of McKibben pneumatic artificial muscle systems," in 2011 IEEE International Conference on Industrial Technology (Auburn, AL, USA: IEEE), 65–70. doi:10.1109/ICIT.2011.5754347

Kalita, B., Leonessa, A., and Dwivedy, S. K. (2022). A review on the development of pneumatic artificial muscle actuators: force model and application. *Actuators* 11, 288. doi:10.3390/act11100288

Kingma, D. P., and Ba, J. (2014). Adam: a method for stochastic optimization. *CoRR abs/1412.6980*. doi:10.48550/arXiv.1412.6980

Lutter, M., Ritter, C., and Peters, J. (2019). Deep lagrangian networks: using physics as model prior for deep learning. *ArXiv abs/1907.04490*. doi:10.48550/arXiv.1907.04490

Ly, K., Yang, Y., Liu, T., Gao, Q., Guo, Q., and Qiu, X. (2023). "Full parameter fine-tuning for large language models with limited resources," in *Annual meeting of the association for computational linguistics*.

Majumdar, R., Jadhav, V., Deodhar, A., Karande, S. S., Vig, L., and Runkana, V. (2023). Hyperlora for pdes. *ArXiv abs/2308.09290*. doi:10.48550/arXiv.2308.09290

Noritsugu, T., and Tanaka, T. (1997). Application of rubber artificial muscle manipulator as a rehabilitation robot. *IEEE/ASME Trans. Mechatronics* 2, 259–267. doi:10.1109/3516.653050

- Olaby, O., Brun, X., Sesmat, S., Redarce, T., and Bideaux, E. (2005). "Characterization and modeling of a proportional valve for control synthesis," in 6th JFPS International Symposium on Fluid Power (Tsukuba, Japan), 771–776. doi:10.5739/isfp.2005.7712005
- Park, Y.-L., Santos, J., Galloway, K. G., Goldfield, E. C., and Wood, R. J. (2014). "A soft wearable robotic device for active knee motions using flat pneumatic artificial muscles," in 2014 IEEE International Conference on Robotics and Automation (ICRA), 4805–4810. doi:10.1109/ICRA.2014.6907562
- Penwarden, M., Zhe, S., Narayan, A., and Kirby, R. M. (2023). A metalearning approach for physics-informed neural networks (pinns): application to parameterized pdes. *J. Comput. Phys.* 477, 111912. doi:10.1016/j.jcp.2023.111912
- Prantikos, K., Chatzidakis, S., Tsoukalas, L. H., and Heifetz, A. (2023). Physics-informed neural network with transfer learning (TL-PINN) based on domain similarity measure for prediction of nuclear reactor transients. *Sci. Rep.* 13, 16840. doi:10.1038/s41598-023-43325-1
- Raissi, M., Perdikaris, P., and Karniadakis, G. (2017). *Physics informed deep learning (part i): data-driven solutions of nonlinear partial differential equations*. doi:10.48550/arXiv.1711.10561
- Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., et al. (2021). Zero-shot text-to-image generation. In *Proceedings of the 38th international conference on machine learning, proceedings of machine learning research (PMLR)*, M. Meila, and T. Zhang editors. vol. 139, 8821–8831.
- Richer, E., and Hurmuzlu, Y. (2000). A high performance pneumatic force actuator system: part I—nonlinear mathematical model. *J. Dyn. Syst. Meas. Control* 122, 416–425. doi:10.1115/1.1286336
- Tondu, B., and Lopez, P. (2000). Modeling and control of McKibben artificial muscle robot actuators. *IEEE Control Syst.* 20, 15–38. doi:10.1109/37.833638
- Torres, E., and Niepert, M. (2024). "Survey: adaptive physics-informed neural networks," in *Neurips 2024 workshop foundation models for science: progress, opportunities, and challenges*.
- Wang, G., Chalard, R., Cifuentes, J., and Pham, M. (2025a). Learning an inverse thermodynamic model for pneumatic artificial muscles control. *Mechatronics* 110, 103359. doi:10.1016/j.mechatronics.2025.103359
- Wang, G., Chalard, R., Cifuentes, J. A., and Pham, M. T. (2025b). "Physics-informed hybrid modeling of pneumatic artificial muscles," in 2025 IEEE International Conference on Robotics and Automation (ICRA), 3407–3413. doi:10.1109/ICRA55743.2025.11127748
- Wang, L., Chen, S., Jiang, L., Pan, S., Cai, R., Yang, S., et al. (2025c). Parameter-efficient fine-tuning in large language models: a survey of methodologies. *Artif. Intell. Rev.* 58, 227. doi:10.1007/s10462-025-11236-4
- Wang, Y., Bai, J., Eshaghi, M., Anitescu, C., Zhuang, X., Rabczuk, T., et al. (2025d). Transfer learning in physics-informed neural networks: full fine-tuning, lightweight fine-tuning, and low-rank adaptation. *Int. J. Mech. Syst. Dyn.* 5, 212–235. doi:10.1002/msd2.70030
- Yaxi, W., and Xu, Q. (2021). Design and testing of a soft parallel robot based on pneumatic artificial muscles for wrist rehabilitation. *Sci. Rep.* 11. doi:10.1038/s41598-020-80411-0
- Zhou, C., Li, Q., Li, C., Yu, J., Liu, Y., Wang, G., et al. (2024). A comprehensive survey on pretrained foundation models: a history from BERT to ChatGPT. *Int. J. Mach. Learn. Cybern.* 16, 9851–9915. doi:10.1007/s13042-024-02443-6