Local Evolution Decoupling Paradigm for Neural Network Constitutive Models of History-Dependent Materials: Theory, Architecture, and Applications to shape memory polymers

A B S T R A C T

The mechanical response of history-dependent materials depends strongly on loading paths and thermomechanical histories, making neural network constitutive modeling inherently challenging. Most existing approaches emphasize fitting complete historical sequences, resulting in excessive parameters, limited generalization, and an inability to initiate predictions from arbitrary intermediate states or variable time steps. This study theoretically demonstrates that for compressible history-dependent materials (CHDMs), the evolution of internal variables exhibits Markovian behavior, revealing structural redundancy in conventional sequence-based neural networks. Accordingly, a Local Decoupling Evolution Operator Network (LDEONet) is proposed, which explicitly decouples internal variable evolution into direction and intensity components. Shape memory polymers (SMPs) are employed as a representative case study. Experimental results show that LDEONet requires only approximately 10^1 parameters and achieves high accuracy using just 1/175 of the training samples required by long short-term memory networks (LSTM). Crucially, LDEONet enables both path and discretization generalization, allowing predictions from arbitrary intermediate states while maintaining numerical stability under constant, random, sinusoidal, and stepwise discretization schemes. Failure boundaries under cross-modal thermomechanical transitions are also examined. LDEONet thus provides a lightweight, high-generalization paradigm for history-dependent material modeling based on local evolution principles.

Keywords: History-dependent materials, Neural network, Constitutive modeling, Shape memory polymers

1. Introduction

Shape memory polymers (SMPs), as a representative class of history-dependent materials, are capable of recovering from a deformed temporary shape to a pre-programmed permanent shape under external stimuli [1, 2]. Their high recoverable strain, low density, tunable glass transition temperature, favorable biocompatibility, and excellent processability render them promising candidates for cutting-edge applications in soft robotics, biomedical devices,

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

space-deployable structures, and four-dimensional (4D) printing [1, 3, 4]. With ongoing technological advances, these application domains impose increasingly stringent requirements on constitutive models capable of describing material behavior under complex loading paths and thermomechanical histories [5, 6].

The intrinsic shape memory mechanism of SMPs originates from the temperature-driven reversible transition between rubbery and glassy states [7]. At elevated temperatures, the material exhibits entropic elasticity, characterized by high mobility of flexible chain segments. As the temperature decreases below the glass transition temperature, molecular segmental motion becomes frozen, and the material enters a glassy state dominated by energetic elasticity [8]. This transition process manifests as the evolution of internal state variables: the stored strain records and freezes the macroscopic deformation experienced during cooling and is gradually released upon reheating above the transition temperature, thereby enabling shape recovery [9, 10].

Extensive research efforts have been devoted to developing constitutive models that describe such thermomechanical behavior. Tobushi et al. first proposed a linear viscoelastic theory to characterize the shape memory effect of SMPs and investigated the thermomechanical response of polyurethane SMPs [11]. Subsequently, linear viscoelastic theory has been progressively refined and extended. Established constitutive models for SMPs can be broadly categorized into viscoelastic models [11], phase transition models [12], and hybrid formulations [13]. Chen and Lagoudas [10] introduced a transformation gradient to establish the relationship between the deformation gradients of the rubbery and glassy phases. Volk et al. [14, 15] experimentally validated and further advanced the finite deformation constitutive theory for SMPs based on comprehensive shape memory characterization of polyurethane. Although these physics-based models offer theoretical rigor and predictive capability, they often rely on cumbersome derivations and involve a large number of material parameters requiring extensive calibration [13]. With the rapid development of SMPs and their associated composites, an increasing variety of physical stimuli and material hybridization strategies are being employed in practical engineering applications [16], further amplifying the complexity of constitutive modeling.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

Consequently, the thermomechanical behavior of SMPs, which inherently depends on the entire loading history, has emerged as both a focal point and a significant challenge in constitutive modeling [17]. Unlike conventional materials, the stress response of history-dependent materials depends not only on the instantaneous strain but also on the entire preceding loading history [18]. Constitutive descriptions of such materials often necessitate the introduction of numerous internal variables to capture their complex evolution laws, which not only complicates the solution procedure but also leads to low computational efficiency, particularly when confronting nonlinearities such as finite deformations and multi-field coupling [19, 20, 21, 22].

To overcome the limitations inherent in traditional modeling paradigms, neural network-based constitutive models have emerged as a promising alternative for history-dependent materials [23, 24, 25]. Among these, recurrent neural networks (RNNs) and their variants, by virtue of their innate sequence-processing capabilities, can automatically encode historical information through hidden states, thereby accurately capturing material memory effects without explicitly defining internal variables [26, 27, 28]. This characteristic has enabled RNN-based models to demonstrate potential exceeding that of exact solutions to traditional differential equations in modeling time-dependent phenomena such as the long-term creep of synthetic concrete [29, 30, 31, 32]. Currently, modeling frameworks based on RNNs [33], long short-term memory networks (LSTM) [34], and gated recurrent units (GRUs) [35] have become mainstream approaches for developing neural network constitutive models for history-dependent materials. Despite their notable successes, these approaches exhibit certain limitations. First, RNN inputs inherently rely on historical sequence data, rendering them incapable of discretization generalization [36]. Second, RNNs exhibit strong sensitivity to the initial hidden state, requiring simulations to commence from a trained reference state; predictions fail when initiated from arbitrary pre-loaded states [37, 38].

To address the limited generalization capabilities of RNNs in material constitutive modeling, researchers have begun exploring paradigms with enhanced physical interpretability. A prominent direction involves incorporating neural ordinary differential equations (NODEs) into the theoretical framework of material evolution. This approach

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

parameterizes the instantaneous terms of dynamical systems and reconstructs material evolution trajectories through numerical integration [39]. Jones et al. proposed the internal state variable neural ODE (ISV-NODE), which implicitly embeds internal variables within the NODE framework, achieving discretization-invariant predictions for viscoelastic and elastoplastic materials [40]. More recently, He and Semnani introduced the incremental neural controlled differential equation (INCDE), which further incorporates stabilization mechanisms and a piecewise formulation to enhance robustness under complex loading paths [41].

In parallel with the aforementioned paradigms, another line of research focuses on directly embedding thermodynamic principles into network architectures by introducing physics-informed constraints to enhance both generalization capability and interpretability. Examples include thermodynamics-based artificial neural networks (TANNs), which enforce thermodynamic consistency [42, 43], and input-convex neural networks, which constrain the convexity of the free energy to ensure the physical validity and stability of the constitutive model [44, 45].

Deep operator networks (DeepONets) and Fourier neural operators (FNOs), as emerging techniques, have achieved substantial progress in scientific artificial intelligence and operator approximation [46, 47], offering new avenues for constitutive modeling of history-dependent materials. Guo et al. proposed the history-aware neural operator (HANO) based on the FNO framework, which constructs an autoregressive formulation to infer future stress directly from stress-strain data within a short-term historical window [48]. By eliminating hidden state propagation and conditioning predictions on recent historical segments, HANO simultaneously mitigates discretization dependence and initial state sensitivity [48]. As an emerging computational paradigm, HANO exhibits considerable potential for practical applications, yet several issues warrant further investigation. For instance, enhancing the physical interpretability of the model with respect to the underlying thermomechanical mechanisms and improving its generalization to unseen loading conditions remain important directions for future research.

These successful cases provide a foundation for developing neural network constitutive

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

models for SMPs. The defining characteristic distinguishing SMPs from conventional materials is the history dependence introduced by the shape memory effect, implying that a constitutive model must be capable of understanding and representing the complex evolution laws embedded within long-range thermomechanical loading histories. Recognizing such laws typically necessitates relatively complex models, which partly accounts for the comparatively slow progress of neural network approaches in SMP constitutive modeling. To date, early exploratory studies have achieved significant advances in this direction. Deng et al. employed an LSTM network to construct a neural network constitutive model for SMPs, demonstrating the capability of LSTM to capture the nonlinear thermomechanical evolution paths of the material [49]. That model contained approximately 1.4×10^{5} parameters and achieved a mean squared error (MSE) on the order of 1.4×10^{-5} [49]. Diverging from the aforementioned paradigm, Yan et al. developed a physics-driven constitutive model based on physics-informed neural networks (PINNs) combined with experimental data, enabling direct mapping from strain and temperature fields to the stress response [50]. This model contained only 273 parameters and, after 2.2×10^{6} iterations, converged to an MSE on the order of 3.0×10^{-4} [50]. Nevertheless, both the models trained by Deng et al. and Yan et al. are not explicitly designed to support path generalization; consequently, initiating predictions from arbitrary initial path conditions remains challenging, which may limit their direct applicability in certain practical computational scenarios.

A crucial theoretical observation provides a new perspective for this work and prompts a re-examination of the limitations inherent in existing methods. A review of the aforementioned studies reveals that prevailing approaches predominantly focus on employing neural networks to approximate a complex memory system [37, 38, 48], to fit global physical laws [42, 43], or to simulate global recurrent states [39, 40, 41]. Such models rely on latent variables to store path sequence information and are inherently dependent on global path data. It is worth noting that although SMPs exhibit pronounced history dependence during macroscopic evolution, the evolution of the material's internal state variables obeys localized equations. This implies that the complex loading history is already fully encoded within the current internal state. This locality property suggests that the instantaneous evolution of the

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

system depends solely on the current state, rather than on explicit reliance on the historical path. Consequently, the modeling emphasis can be shifted from fitting complex long-range memory effects toward learning local evolution laws, thereby mitigating model dependence on specific paths and discretization schemes while simultaneously reducing model complexity.

Following this line of reasoning, this paper first revisits the internal variable evolution equations for SMPs, subsequently elucidates the Markovian property inherent in these equations, and demonstrates the structural redundancy of conventional sequence-based models in predicting compressible history-dependent materials (CHDMs). Building upon this theoretical foundation, a Local Decoupling Evolution Operator Network (LDEONet) is proposed, which theoretically achieves both path generalization and discretization generalization capabilities. By emulating the local evolution laws of internal variables, LDEONet explicitly decouples the evolution direction from the evolution intensity, thereby reducing both sample requirements and model complexity. Subsequently, the performance of the proposed model is trained and evaluated using shape memory polymers as a representative case study. Finally, a comprehensive suite of generalization experiments is designed to validate the generalization capability of LDEONet.

2. Internal Variable Evolution and Its Markovian Property

2.1 Internal Variable Evolution Model

Shape memory polymers, as a representative class of history-dependent materials, exhibit macroscopic mechanical responses that depend not only on instantaneous loading but also on the entire history of coupled thermomechanical loading. The macroscopic shape memory effect originates microscopically from the cumulative manifestation of the time-dependent evolution of the material's internal structure, i.e., the hindered motion and thermally activated processes of polymer chain segments across various energy barriers. To describe the deformation process, the macroscopic total strain tensor of SMPs, denoted by $\boldsymbol{\varepsilon}$, can be decomposed into the sum of the elastic strain tensor $\boldsymbol{\varepsilon}_e$ induced by the instantaneous stress, the thermal expansion strain tensor $\boldsymbol{\varepsilon}_T$ induced by temperature variation, and the stored

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

strain tensor $\boldsymbol{\varepsilon}_s$ characterizing the shape memory effect:

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_e + \boldsymbol{\varepsilon}_T + \boldsymbol{\varepsilon}_s \tag{2}$$

Here, the thermal strain $\boldsymbol{\varepsilon}_T$ is obtained by integrating the coefficient of thermal expansion α with respect to temperature T:

$$\boldsymbol{\varepsilon}_T = \int_{T_0}^{T} \alpha \, \mathrm{d}T \tag{3}$$

Based on the series model derived from phenomenological theory, the deformation of SMPs can be regarded as the combined response of the glassy phase and the rubbery phase subjected to an identical stress state [51]. Accordingly, the equivalent compliance tensor of the material $\boldsymbol{S}(T)$, can be expressed in terms of the compliance tensor of the glassy phase $\boldsymbol{S}_g$, and that of the rubbery phase $\boldsymbol{S}_r$, as follows:

$$\boldsymbol{S}(T) = \phi_g \boldsymbol{S}_g + (1 - \phi_g) \boldsymbol{S}_r \tag{4}$$

where ϕ_g denotes the volume fraction of the glassy phase. The equivalent stiffness tensor $\boldsymbol{C}$ and the corresponding constitutive relation can be formulated as:

$$\boldsymbol{\sigma} = \boldsymbol{C} : \boldsymbol{\varepsilon}_e = (\phi_g \boldsymbol{S}_g + (1 - \phi_g) \boldsymbol{S}_r)^{-1} : \boldsymbol{\varepsilon}_e \tag{5}$$

The physical admissibility of the foregoing kinematic description and constitutive model must be rigorously constrained by the laws of thermodynamics. Based on the free energy function and the Clausius-Duhem inequality, Liu et al. derived a thermodynamically consistent local evolution law for the stored strain $\boldsymbol{\varepsilon}_s$ [51]:

$$\frac{\mathrm{d}\boldsymbol{\varepsilon}_s}{\mathrm{d}T} = \boldsymbol{S}_e(T) : \boldsymbol{\sigma}(T) \frac{\mathrm{d}\phi_g}{\mathrm{d}T} \tag{6}$$

where $\boldsymbol{S}_e(T) = 3NkT\boldsymbol{I}$ denotes the entropic compliance tensor, with N being the crosslink density, k the Boltzmann constant, and $\boldsymbol{I}$ the fourth-order identity tensor. To effectively utilize the experimental data reported by Li, Volk, and co-workers [14, 15, 16], the present study approximates the entropic compliance tensor by the experimentally measured compliance tensor of the rubbery phase. Consequently, the evolution law for the internal variable $\boldsymbol{\varepsilon}_s$ adopted herein is expressed as:

$$\frac{\mathrm{d}\boldsymbol{\varepsilon}_s}{\mathrm{d}T} = \boldsymbol{S}_r(T) : \boldsymbol{\sigma}(T) \frac{\mathrm{d}\phi_g}{\mathrm{d}T} \tag{7}$$

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

2.2 Markovian Property of the Evolution

Although the coupled thermomechanical evolution of SMPs is history-dependent, the evolution law governing the internal variables is inherently history-independent. Substituting **Equation 4** into **Equation 6** to eliminate the stress tensor $\boldsymbol{\sigma}$ yields:

$$\frac{\mathrm{d}\boldsymbol{\varepsilon}_s}{\mathrm{d}T} = \boldsymbol{S}_r : \left(\phi_g \boldsymbol{S}_g + (1-\phi_g)\boldsymbol{S}_r\right)^{-1} : \boldsymbol{\varepsilon}_e \frac{\mathrm{d}\phi_g}{\mathrm{d}T} \tag{8}$$

Treating the glassy phase compliance $\boldsymbol{S}_g$ and the rubbery phase compliance $\boldsymbol{S}_r$ as constants, the mapping relation $\mathscr{F}$ governing the internal variable evolution can be summarized as follows:

$$\frac{\mathrm{d}\boldsymbol{\varepsilon}_s}{\mathrm{d}T} = \mathscr{F}\left(\boldsymbol{\varepsilon}_e, \phi_g, \frac{\mathrm{d}\phi_g}{\mathrm{d}T}\right) \tag{9}$$

Examination of the variable composition in **Equation 8** reveals that the mapping relation $\mathscr{F}$ is defined exclusively by the instantaneous state of the material, namely $\boldsymbol{\varepsilon}_e$, ϕ_g, and $\mathrm{d}\phi_g/\mathrm{d}T$. A system whose state update depends solely on the current instantaneous values, irrespective of the past history by which that state was attained, is said to satisfy the Markovian property. As illustrated in **Figure 1**, this observation does not negate the history-dependent nature of the evolution of shape memory polymers; rather, it indicates that the memory effects accumulated throughout the loading history are already completely and physically encoded within the current internal state variables, obviating the need to explicitly trace the complete loading history.

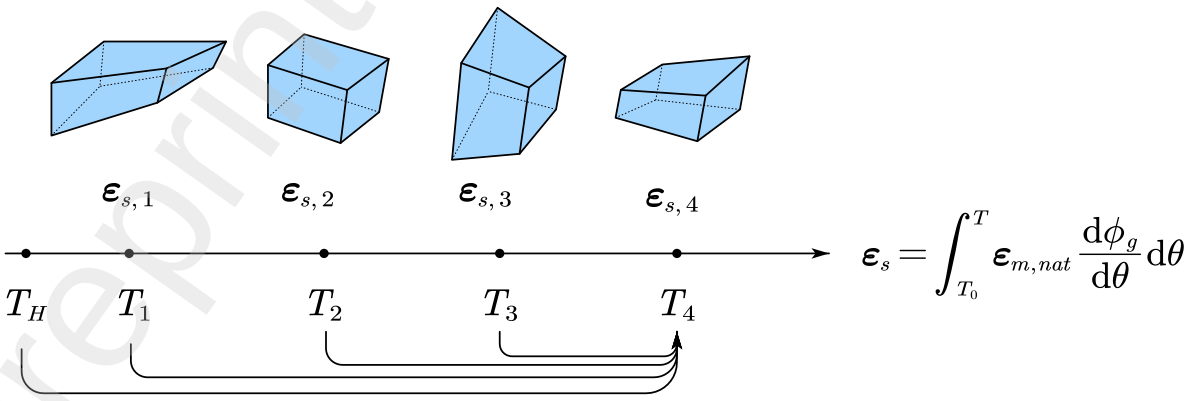


Figure 1: Evolution process of the internal variable

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

Assuming that the evolution of the internal variable constitutes a dynamical system $\mathcal{X}(T)$ defined over time, and in accordance with the mapping relation for the internal variable, the dynamical system can be expressed as:

$$\mathcal{X}(T) = \mathscr{F}(T) \tag{10}$$

According to the fundamental theorem of calculus, the future state of the internal variable can be obtained by integrating the current dynamical system:

$$\boldsymbol{\varepsilon}_s(T+\Delta T) = \boldsymbol{\varepsilon}_s(T) + \int_T^{T+\Delta T} \mathcal{X}(\tau)\,\mathrm{d}\tau \tag{11}$$

If the dynamical system possesses the Markovian property, the evolution mapping relation $\mathscr{F}$ must constitute a purely local algebraic function, devoid of any internal memory integral kernels or hereditary functions that depend on the historical time $t \in (-\infty, \tau)$. From the perspective of dynamical systems theory, this implies that the evolved state of the system at time $t+\Delta t$, denoted by $\mathcal{X}_{t+\Delta t}$, can be uniquely determined by the instantaneous characteristics $\mathcal{X}_t$ at time t and the variation occurring within that time increment. The future evolution trajectory of the system is independent of the specific historical loading path $\{\mathcal{X}_{t-\Delta t}, \ldots, \mathcal{X}_0\}$ traversed prior to reaching the current state $\mathcal{X}_t$. Evidently, the mapping relation $\mathscr{F}$ governing the internal variable evolution of shape memory polymers satisfies this property and therefore exhibits the Markovian property. In fact, any material satisfying this property is Markovian in nature. The relevant exposition can be extended to a broader class of material systems termed Compressible History-Dependent Materials (CHDMs), whose defining characteristic is that the material response can be described in closed form using a finite-dimensional set of state variables.

Within the framework of material constitutive theory, many path-dependent materials are often described using integral-type memory models, the general form of which is given by:

$$\sigma(t) = \int_0^t K(t-s)\,\mathcal{F}(\varepsilon(s), T(s))\,\mathrm{d}s \tag{12}$$

where $K(t)$ denotes the memory kernel function, which characterizes the influence of the historical loading on the current response. Models of this form exhibit an explicit

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

dependence on the complete historical path and thus constitute a typical non-local temporal description of material evolution.

However, for a class of memory kernel functions possessing a finite spectral expansion. For instance, those representable as a superposition of exponential functions:

$$K(t) \approx \sum_{i=1}^{N} a_i e^{-t/\tau_i} \tag{13}$$

the corresponding hereditary integral representation can be equivalently transformed into a set of evolution equations for a finite number of internal variables:

$$\dot{z}_i = \frac{1}{\tau_i}\left(\mathcal{F}(\varepsilon, T) - z_i\right), \quad i = 1, \ldots, N \tag{14}$$

This transformation recasts the model, which originally depended on the complete loading history, into a local dynamical system that is closed in the state space. Under these circumstances, the evolution of the material system satisfies the Markovian property: its future response depends solely on the current state variables and no longer requires explicit reference to the historical loading path.

2.3 Limitations of Existing Models

Path-dependent, data-driven models, represented by RNNs, LSTMs, and HANO, possess inherent advantages in capturing the history dependence of materials by virtue of their specialized memory mechanisms. Through deep learning on large volumes of data, these models can circumvent cumbersome constitutive derivations and implicitly fit the nonlinear coupled thermomechanical constitutive mapping relations. However, models trained on sequential data tend to be influenced, from an architectural standpoint, by the path dependence and discretization dependence present in the training data. Consequently, achieving robust path and discretization generalization with such architectures remains challenging. HANO systematically summarized the limitations of models such as RNNs and LSTMs, establishing a predictive model that relies on sequential windows. This approach alleviates the dependence on global sequence data and achieves a certain degree of path generalization capability [48].

To overcome the intrinsic deficiencies of conventional models, Neural Ordinary

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

Differential Equations (NODEs) and their variant, Internal State Variable Neural ODEs (ISV-NODEs), embed neural networks into the right-hand side of differential equations, thereby incorporating the instantaneous evolution of internal variables within a numerical integration framework [39, 40, 41]. This formulation approximates the entire thermomechanical loading history and thus attains discretization generalization capability. Thermodynamics-based Artificial Neural Networks (TANNs) integrate thermodynamic scalar potentials and constraint equations into the model architecture, significantly enhancing the thermodynamic consistency of the implicit internal variables [42]. Physics-Informed Neural Networks (PINNs) employ evolution equations that satisfy thermodynamic consistency as loss functions to directly train constitutive relations for SMPs [50]. Neural network models such as PINNs, NODEs, and TANNs perform notably well in scenarios characterized by relatively sparse data availability. By leveraging physical constraints—including the laws of thermodynamics, free energy dissipation mechanisms, or internal variable evolution equations—these models are capable of rigorously preserving the physical consistency of their predictions. Nevertheless, despite adopting local evolution laws and explicit physical models in their mathematical formulation, the training process of these models typically depends on full historical path data. This implicit reliance on global training data implies that the models remain fundamentally constrained by the loading history.

The foregoing exposition concerning the locality of internal variable evolution has demonstrated that the evolution of internal variables in CHDMs satisfies the Markovian property. This implies that machine learning models developed for this class of material systems need not model the historical path data explicitly and are inherently endowed with both path generalization and discretization generalization capabilities. In fact, predicting the subsequent evolution of such materials does not require revisiting the entire sequence of past events that the material has undergone; rather, it suffices to concentrate solely on the instantaneous physical state. However, existing models frequently adopt strategies that contradict this fundamental physical nature [49, 50], thereby resulting in structural redundancy within the model architecture.

3. Model, Sample Generation, and Training

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

3.1 Local Decoupling Evolution Operator Network

In the field of operator learning, DeepONet achieves universal approximation of nonlinear operators through a decoupled dual-network architecture comprising a branch network and a trunk network [46]. This architectural philosophy of decomposing complex laws into two independent information streams provides important inspiration for the present work. By abstracting the complex evolution law of internal variables as an operator mapping problem on the tangent bundle of the state space, and by explicitly decoupling the evolution direction D from the evolution intensity M, the internal variable evolution equation for SMPs can be formulated as:

$$\frac{\mathrm{d}\boldsymbol{\varepsilon}_s}{\mathrm{d}T} = \mathscr{F}\left(\boldsymbol{\varepsilon}_e, \phi_g, \frac{\mathrm{d}\phi_g}{\mathrm{d}T}\right) = D(\mathbf{s}) \bowtie M(\mathbf{t}) = D(\boldsymbol{\varepsilon}_e, \phi_g) \bowtie M\left(\frac{\mathrm{d}\phi_g}{\mathrm{d}T}\right) \tag{15}$$

As illustrated in **Figure 2**, from the perspective of differential geometry, if the ensemble of all possible physical states of the material (e.g., strain, phase fraction) is defined as a high-dimensional state manifold $\mathscr{L}$, then the loading history of the material can be represented as a continuous curve on the manifold $\mathscr{L}$. Based on the local evolution nature elucidated previously, the evolution direction $\partial_t \boldsymbol{\Gamma}$ of this curve at any point should be entirely determined by the location of that point on the manifold, the instantaneous driving intensity applied at that point, and the coupling mode. Accordingly, the Local Decoupling Evolution Operator Network (LDEONet) is proposed and expressed as:

$$\partial_T \boldsymbol{\Gamma} = \mathrm{LDEONet}(\boldsymbol{p}, \boldsymbol{q}) = D(\boldsymbol{p}) \bowtie M(\boldsymbol{q}) \tag{16}$$

where $\boldsymbol{p}$ and $\boldsymbol{q}$ denote the model prediction parameter vectors. Based on the obtained instantaneous evolution rate, LDEONet is embedded within a temporal integration scheme, enabling the determination of the stored strain at the subsequent temperature increment ΔT as follows:

$$\boldsymbol{\varepsilon}_s(T + \Delta T) = \boldsymbol{\varepsilon}_s(T) + \int_T^{T+\Delta T} \partial_T \boldsymbol{\Gamma}(\tau)\, \mathrm{d}\tau \tag{17}$$

Subsequently, based on the stored strain, elastic strain, and thermal strain at the next step, the corresponding stress can be determined as:

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

$$\boldsymbol{\sigma}(T+\Delta T)=\boldsymbol{C}:(\boldsymbol{\varepsilon}(T+\Delta T)-\boldsymbol{\varepsilon}_s(T+\Delta T)-\boldsymbol{\varepsilon}_T(T+\Delta T)) \tag{18}$$

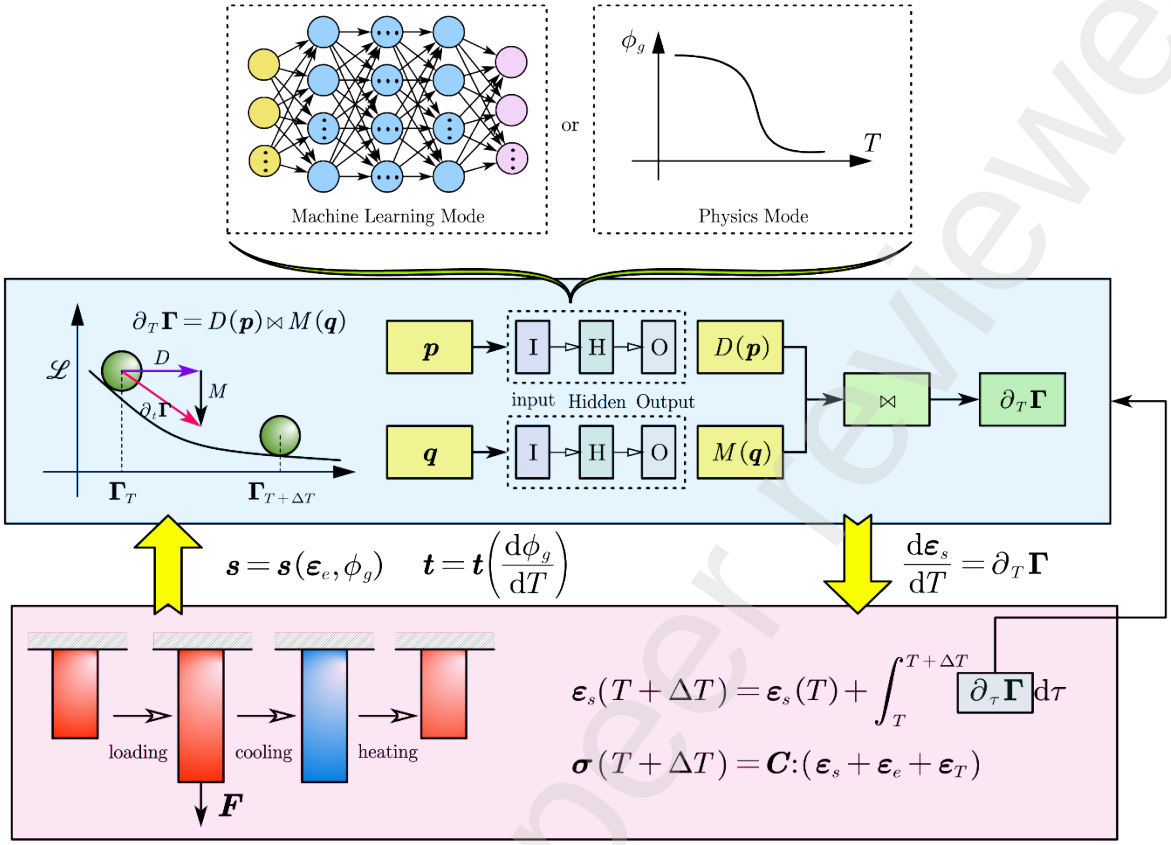


Figure 2: Schematic illustration of the LDEONet model architecture

Figure 2 illustrates the model architecture and the underlying modeling philosophy of LDEONet, namely, utilizing a machine learning model to replace the core evolution model within a complex physical process to achieve more generalizable and fundamental predictions. LDEONet endows the purely data-driven model for the internal variable evolution law of shape memory polymers with a distinct differential geometric interpretation. Through the decoupling of evolution direction and evolution intensity, the physical interpretability of the model is substantially enhanced. Consequently, LDEONet constitutes a neural network model with an inherently physics-structured architecture, rather than merely embedding physical models into the loss function. Furthermore, this architectural-level decoupling imparts modular characteristics to the model: both the evolution direction and the evolution intensity can be directly replaced or augmented by specific physical models according to research requirements. This implies that LDEONet can transform established

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

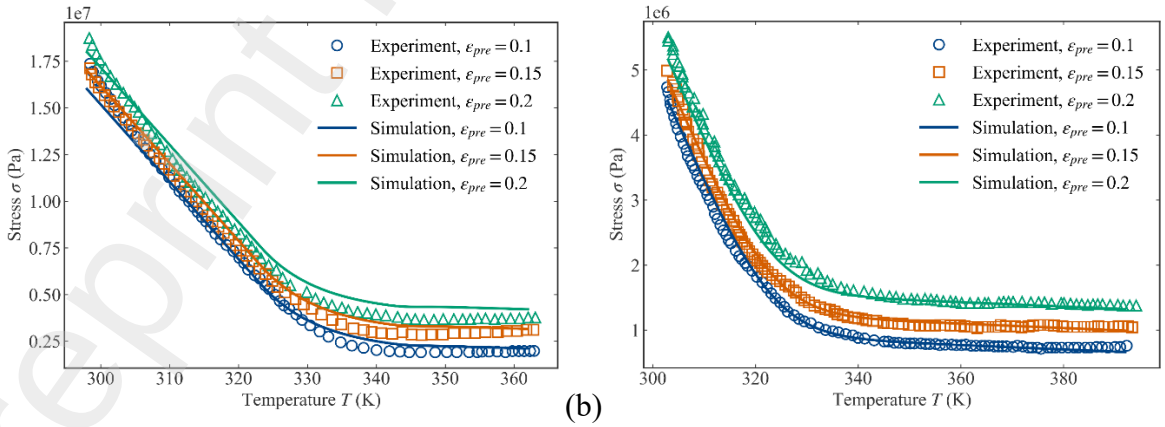
knowledge into pluggable model modules, thereby unlocking the generalization potential of data-driven approaches while preserving physical determinism.

3.2 Sample Generation

As a representative class of compressible history-dependent materials, SMPs exhibit complex spatiotemporal evolution characteristics, making them an ideal testbed for evaluating the path generalization and discretization generalization capabilities of LDEONet. Within the shape memory cycle, the cooling and fixing stage essentially constitutes a nonlinear evolution and storage process of the material's internal state variables, thereby epitomizing the history-dependent nature of the material. Consequently, the uniaxial cooling and fixing process is selected as the core research scenario in this study to validate the model's capacity to capture complex state evolution. The finite difference method is employed to numerically simulate the stored strain evolution equation of SMPs, thereby generating a one-dimensional sample database. The numerical scheme is expressed as:

$$\frac{\boldsymbol{\varepsilon}_s^{n+1} - \boldsymbol{\varepsilon}_s^{n}}{\Delta T} = \mathscr{F}^n \tag{19}$$

where $\boldsymbol{\varepsilon}_s^{n+1}$ denotes the stored strain tensor at the subsequent temperature increment. As a purely data-driven framework, LDEONet can achieve higher predictive accuracy upon acquiring more precise data. By incorporating the experimentally measured glassy phase fraction curve, temperature-stress data under arbitrary pre-strain conditions can be obtained.



(a) (b)

Figure 3: Comparison between numerical simulation results and experiments: (a) Comparison with experiments by Volk et al.; (b) Comparison with experiments by Li et al.

Figure 3 presents the temperature-stress variation trends derived from numerical

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

simulation and experimental data. The two trends are largely consistent, with certain discrepancies observed within the temperature interval characterized by more pronounced phase transition, although the degree of deviation remains acceptable. Benefiting from the explicit decoupling of the local evolution law, LDEONet can leverage physical models to reduce model complexity. Specifically, the relationship between the glassy phase fraction and temperature $\phi_g = \phi_g(T)$, has been experimentally determined, from which $\mathrm{d}\phi_g/\mathrm{d}T$ can be derived. Consequently, the evolution intensity M can be replaced by a physical model, and LDEONet is specialized as:

$$\frac{\mathrm{d}\boldsymbol{\varepsilon}_s}{\mathrm{d}T} = \mathrm{LDEONet}(\boldsymbol{p},\boldsymbol{q}) = D(\boldsymbol{p}) \bowtie M(\boldsymbol{q}) = D(\boldsymbol{\varepsilon}_e,\phi_g)\frac{\mathrm{d}\phi_g}{\mathrm{d}T} \tag{20}$$

Therefore, during the practical implementation, only the evolution direction D of LDEONet requires training. Considering the differences in orders of magnitude among the various data, the sample data are normalized via scaling. By performing batch sample expansion with a pre-strain increment of 0.01 within the range $\varepsilon_{pre} \in [0, 0.2]$, a total of 21 cooling and fixing process samples is acquired. Furthermore, the dataset is partitioned according to a ratio of 8:1:1 for the training, validation, and testing of LDEONet, respectively. Owing to the locality property of LDEONet, the continuous evolution path can be decomposed into discrete state points. Each discrete point from the 21 process samples is thereby transformed into a state-point sample, ultimately yielding 1440 training samples, 180 validation samples, and 180 test samples. Finally, the training samples are randomly shuffled to eliminate any sequence-induced bias in the sample data, thereby compelling the model to learn a path-independent mapping mechanism.

3.3 Training

To train and test LDEONet, an internally developed materials artificial intelligence platform was utilized. This platform was developed based on reverse-mode automatic differentiation and tape-recording techniques. The Mean Squared Error (MSE) function is adopted as the loss function to quantify the discrepancy between predicted and actual values:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{21}$$

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

Model parameter optimization is performed using the Adam algorithm, with an initial learning rate set to 10^{-3} and a decay rate set to 10^{-4}. To enhance training accuracy and mitigate oscillations, a piecewise learning rate decay strategy is employed, wherein the learning rate is reduced by a factor of 1/10 every 5000 iterations. Considering the drastic numerical variation of the glassy phase fraction during the phase transition process, conventional saturating activation functions are prone to inducing vanishing gradients. Hence, the Rectified Linear Unit (ReLU) is employed as the hidden layer activation function, leveraging its unbounded positive region to prevent gradient vanishing, and is combined with a linear output layer to ensure convergence of the model on the internal variable evolution law.

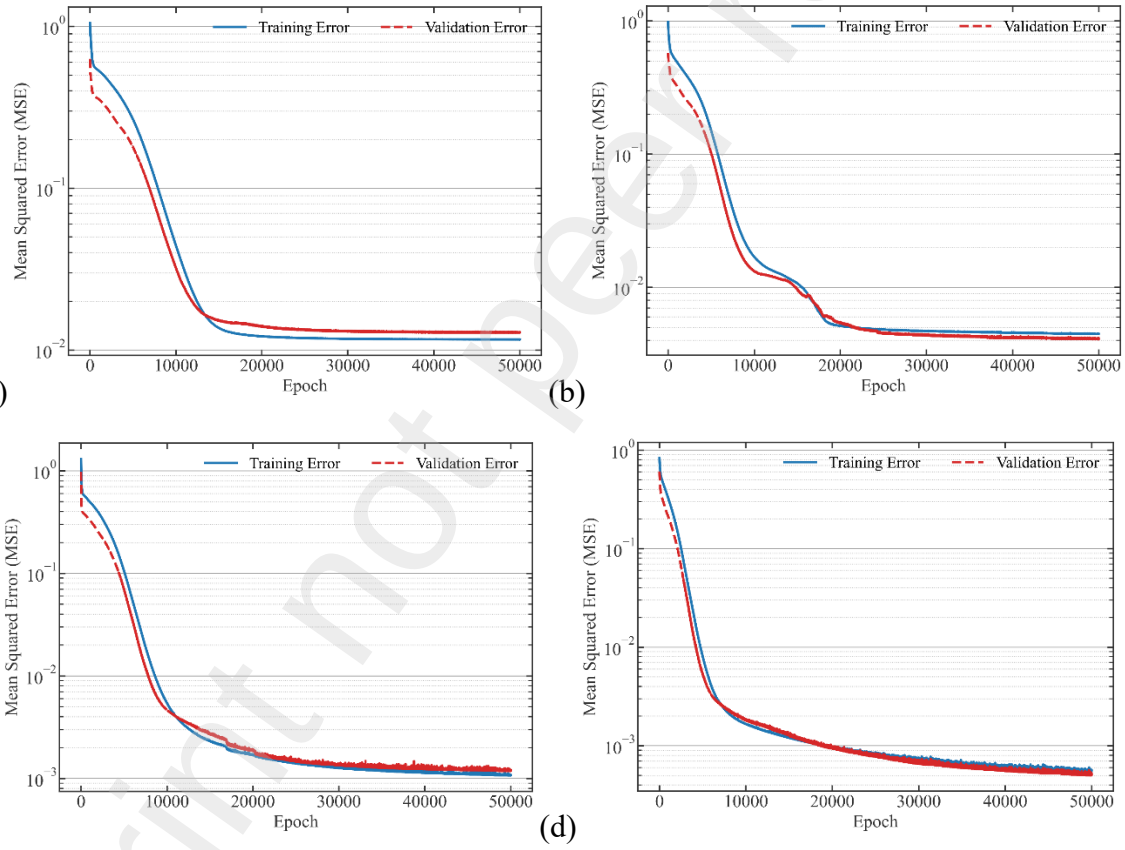


Figure 4: Loss function convergence under different configurations: (a) LDEONet-A1; (b) LDEONet-A2; (c) LDEONet-A3; (d) LDEONet-A4

Following the decoupling of the local evolution law, the complex constitutive model is reduced to a simple local evolution direction model, possessing only two inputs and a single output. Consequently, a single hidden layer is adopted, and various configurations concerning the number of hidden layer neurons are tested to examine the parameter requirements of

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

LDEONet, as presented in **Table 1**. As illustrated in **Figure 4**, based on the training results, the LDEONet-A1 model comprises a total of 26 parameters; however, its MSE is on the order of 10^{-2}, with an R^2 value (coefficient of determination) of 0.9833. For the training results of LDEONet-A2, LDEONet-A3, and LDEONet-A4, the R^2 values all exceed 0.999. Among these, LDEONet-A4 attains an MSE on the order of 10^{-4}. Thus, LDEONet-A1 appears to possess a sufficient number of parameters and has learned the correct internal variable evolution law.

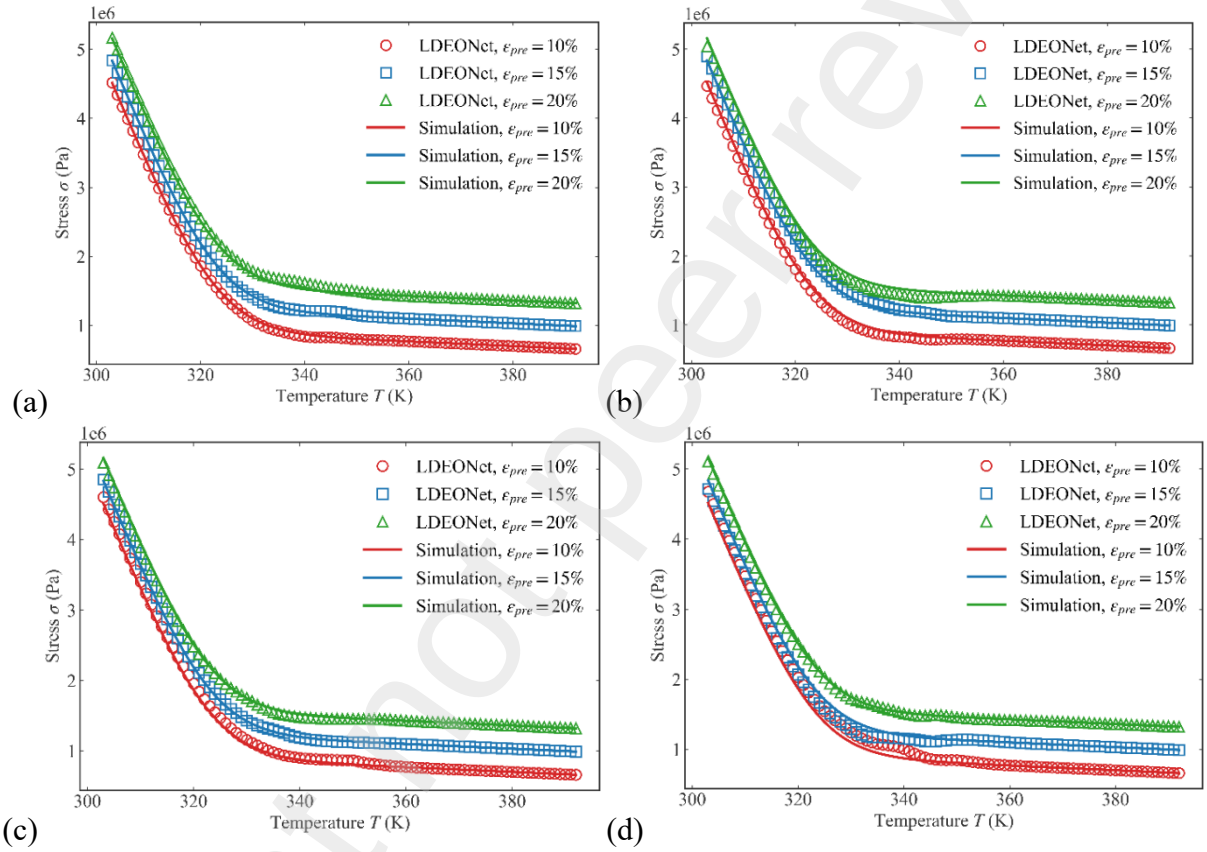


Figure 5: Comparison of predicted stress with experimental data from Li et al. under various configurations: (a) LDEONet-A1; (b) LDEONet-A2; (c) LDEONet-A3; (d) LDEONet-A4

To intuitively compare model performance and evaluate the practical predictive capability for SMPs, a prediction performance test experiment is designed based on the shape memory polymer model. Using experimentally validated numerical results as the benchmark, LDEONet-A1, LDEONet-A2, LDEONet-A3, and LDEONet-A4 are each substituted for the internal variable evolution equation component within the numerical simulation, and predictions are initiated. The prediction starting point is $T_0 = 393\ \mathrm{K}$, with a step size of

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

$\Delta T = -1$ K, and the end point is $T_1 = 303$ K. The pre-strain values considered are $\varepsilon_{pre} = 0.1, 0.15, 0.2$, respectively. As shown in **Figure 5**, although LDEONet-A1 achieves an R^2 value of 0.9833, the actual prediction results demonstrate satisfactory performance only for the condition $\varepsilon_{pre} = 0.2$, with predictions for other conditions deviating from the sample curves. The performance test confirms that, in practical application, LDEONet-A1, LDEONet-A2, LDEONet-A3, and LDEONet-A4 exhibit sufficient predictive capability, among which LDEONet-A4 achieves the most favorable predictive performance.

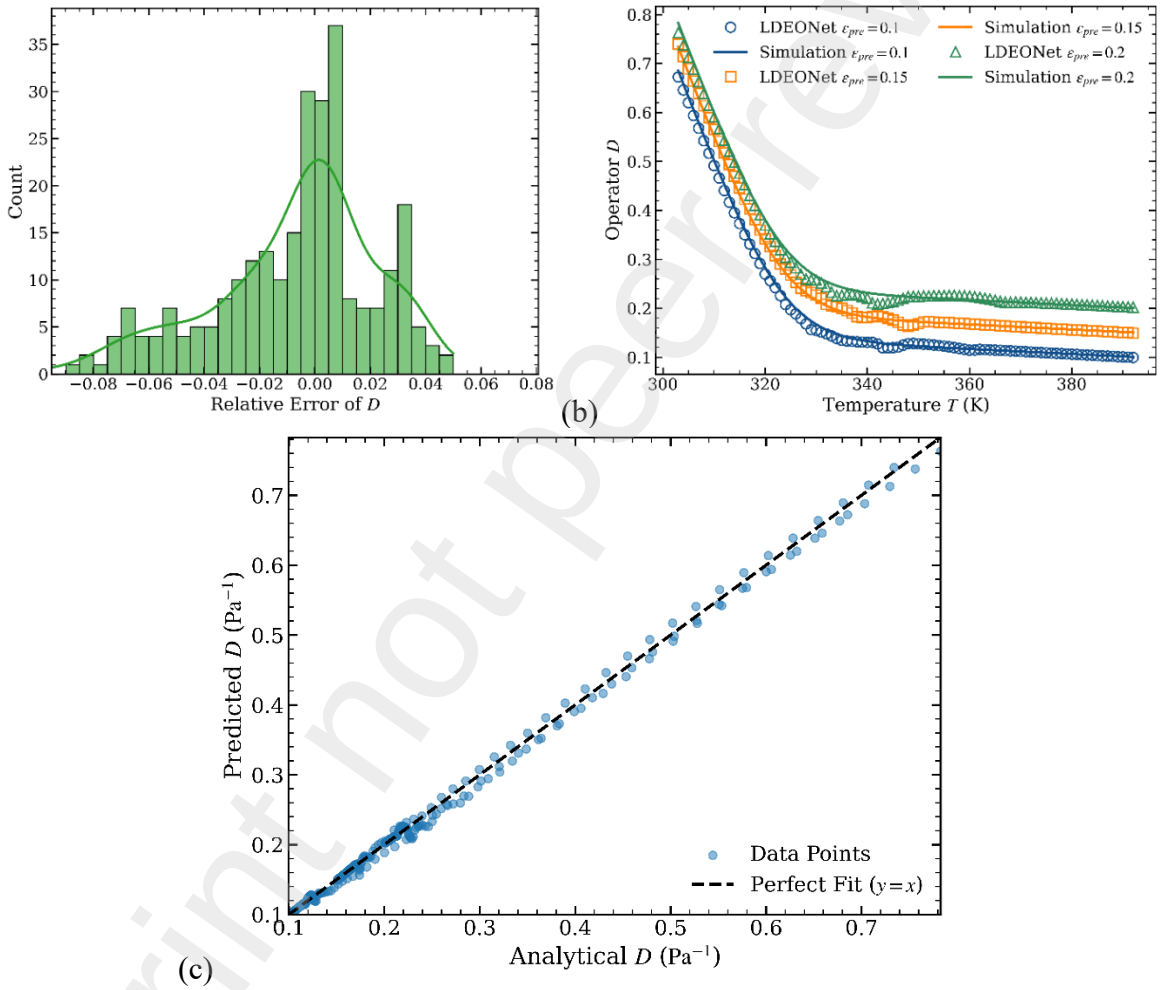


Figure 6: Error analysis of the LDEONet-A4 model: (a) Error distribution; (b) Direction operator D; (c) Parity plot

To analyze the reliability of the LDEONet-A4 model, the error distribution, the learned direction neural operator D, and a parity plot are presented in **Figure 6**. In **Figure 6(a)**, the error distribution of LDEONet-A4 is relatively concentrated, indicating that the predictive deviation of the model is acceptable. **Figure 6(b)** reveals that the direction neural operator D

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

is largely consistent with the sample data, albeit with minor oscillations observed in the vicinity of the glass transition temperature ($T_g = 343\ \mathrm{K}$). These oscillations are attributable to the strongly nonlinear evolution characteristics within this interval, and their magnitude remains acceptable. **Figure 6(c)** presents the parity plot, demonstrating a good degree of fit between the predictions and the sample values. Accordingly, LDEONet-A4 is designated as the primary test subject in subsequent model evaluations. It is noteworthy that the LDEONet-A4 model achieves relatively accurate predictive performance with merely 151 parameters, successfully learning the internal variable evolution law with a modest parameter count.

4. Discussion on Generalization Performance

In this section, a series of experiments are designed to evaluate the path generalization performance, discretization generalization performance, and mode generalization performance of LDEONet. The generalization capability of LDEONet-A4 is systematically tested. Subsequently, LDEONet-A4 is compared with existing models to elucidate its generalizability, limitations, and applicable scope.

4.1 Path Generalization Performance

In practical numerical simulations and experimental investigations, neural network models are required to possess predictive capability from arbitrary path points; therefore, path generalization constitutes one of the critical generalization capabilities of the model. As illustrated in **Figure 7**, conventional models such as RNNs and LSTMs rely on the matching of global path data to produce reasonable predictions of the constitutive relations of history-dependent materials. This implies that the model can only generate correct predictions when the current loading conditions are entirely consistent with its stored global memory. This limitation fundamentally originates from the recurrent architecture inherent to RNNs and their variants. Consequently, models such as RNNs encounter considerable difficulty in achieving genuine path generalization for materials. Certain more advanced models, such as HANO, achieve enhanced path generalization by matching historical path windows to predict the material system at the subsequent time step. In contrast to conventional models, LDEONet

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

focuses exclusively on learning the local evolution law of internal variables. Owing to the Markovian property of the internal variable evolution law of SMPs, LDEONet obviates the need for window-based memory or global memory and is inherently endowed with the capacity to initiate predictions from arbitrary states.

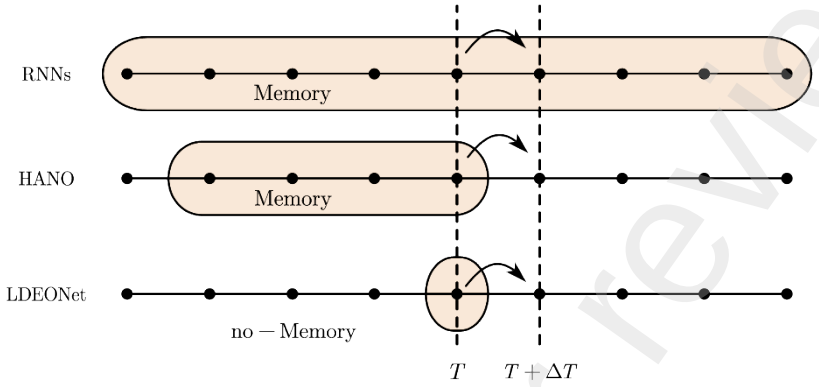


Figure 7: Dependence on historical memory

To validate the path generalization capability of the LDEONet architecture, the LDEONet-A4 model is employed as the object of analysis, and the phase transition states of SMPs are designated as critical nodes in the coupled thermomechanical evolution process. A path generalization performance experiment is accordingly designed. The experiment utilizes the sample database as a benchmark to examine the model's ability to initiate predictions from arbitrary path points and reasonably forecast the subsequent evolution process. Specifically, the starting points of the experiment are set at different glassy phase volume fractions ($\phi_g = 0.2, 0.4, 0.6, 0.8$), the prediction step size is taken as $\Delta T = -1\ \mathrm{K}$, the terminal temperature is $T_1 = 303\ \mathrm{K}$, and the pre-strain conditions remain consistent with those described previously.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

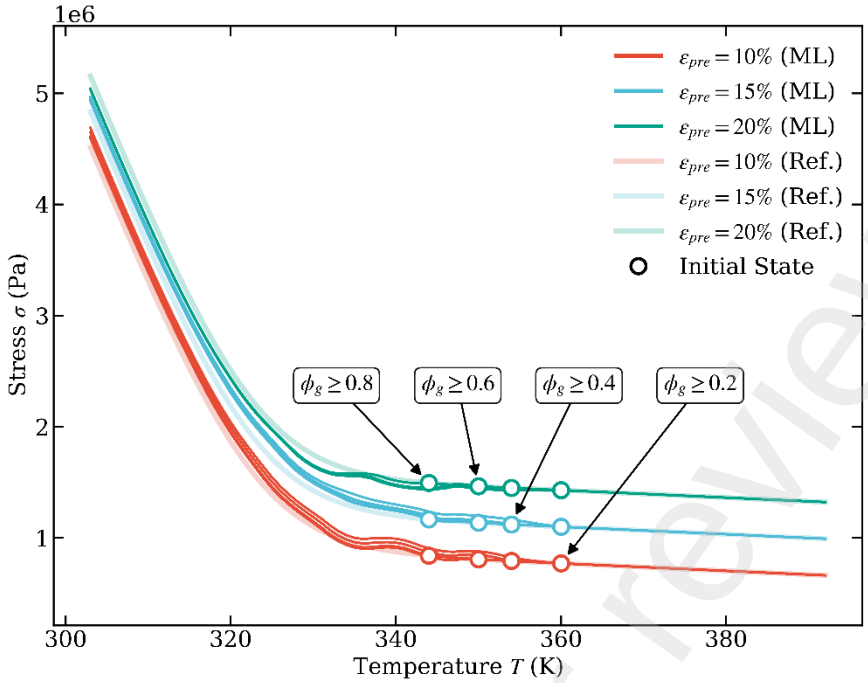


Figure 8: Path generalization performance

As shown in **Figure 8**, the LDEONet-A4 model successfully initiates predictions from starting points corresponding to different glassy phase volume fractions ($\phi_g = 0.2, 0.4, 0.6, 0.8$), and the subsequent prediction trajectories are largely consistent with the benchmark samples. Despite having access solely to instantaneous state point data and being unable to trace the heating and initial loading history of the material, the model successfully executes predictions under all four phase transition states representing distinct microstructural proportions. Observing the evolution curves reveals that, whether in the initial stage dominated by the rubbery phase at elevated temperatures ($\phi_g \geq 0.2$) or in the final stage characterized by a higher fraction of the glassy phase at lower temperatures ($\phi_g \geq 0.8$), the stress evolution trajectories predicted by the LDEONet-A4 model remain consistent with the benchmark data. No numerical oscillations or drift phenomena attributable to the absence of historical sequences are observed. The slight oscillations appearing after prediction initiation in the figure are consistent with the earlier analysis, exert a limited influence on subsequent predictions, and the results are deemed acceptable. Ultimately, the predictions for three pre-strain conditions across four critical path starting points substantiate that the LDEONet-A4 model possesses robust path generalization capability. Furthermore, the

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

prediction results corroborate the rationale behind decoupling the evolution direction from the evolution intensity within the LDEONet architecture, demonstrating that training solely the evolution direction enables the LDEONet-A4 model to learn a robust internal variable evolution law.

4.2 Discretization Generalization Performance

In the core application scenarios of neural network models, discretization generalization capability is a property of equal importance to path independence. In finite element analysis, in addition to the diversity of initial conditions, the diversity of temporal discretization schemes cannot be overlooked. Factors such as heating rate and time step size are critical elements influencing both efficiency and convergence, and they are often adjusted dynamically during the solution process. Owing to their inherent architectural limitations, conventional models such as RNNs exhibit a strong coupling between the weight updates of their recurrent units and the sampling interval of the training data during the training phase. When the step size varies during the prediction phase, the local feature distribution of the model input sequence deviates from the training distribution, leading to amplified cumulative errors. To validate the discretization invariance of LDEONet, this subsection designs a discretization generalization performance experiment based on various variable temperature step size schemes, testing the sensitivity of the LDEONet-A4 model to discretization. During practical implementation, except for the discretization conditions controlled by the experimental scheme, all other experimental conditions remain identical to those described previously. Four non-stationary discretization schemes are specifically devised, all of which ensure $\Delta T < 0$ to guarantee the stability of the cooling and fixing process, as illustrated in **Figure 9**.

Figure 9(a) presents the constant baseline discretization scheme, employing a uniform step size of $\Delta T = -1.0\ \mathrm{K}$, and serves as the reference case for the other discretization schemes. Under this condition, the LDEONet-A4 model achieves favorable predictions of the coupled thermomechanical process during cooling and fixing. **Figure 9(b)** illustrates the random noise discretization scheme, wherein the step size obeys a uniform random

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

distribution $\Delta T \sim \mathcal{U}(-2.0, -0.1)\,\mathrm{K}$, simulating high-frequency perturbations in the temperature control system. The results demonstrate that the LDEONet-A4 model is virtually unaffected by high-frequency perturbations and continues to yield satisfactory predictions. **Figure 9(c)** depicts the sinusoidal fluctuation discretization scheme, where the absolute step size varies periodically in a sinusoidal manner, $\Delta T = -|1.5\sin(\cdot)|$, testing numerical stability under continuously fluctuating step sizes. The LDEONet-A4 model maintains excellent predictive capability and remains largely impervious to continuous fluctuations. **Figure 9(d)** shows the stepwise variation discretization scheme, wherein the step size discontinuously jumps between $-0.1\,\mathrm{K}$ and $-2.0\,\mathrm{K}$, examining the transient response of the model to abrupt step size changes. The LDEONet-A4 model still performs favorable predictions, with virtually no discernible impact of the jumps on the prediction results. This indicates that the model possesses strong robustness against step size variations and is capable of maintaining consistent constitutive responses under these extreme tests.

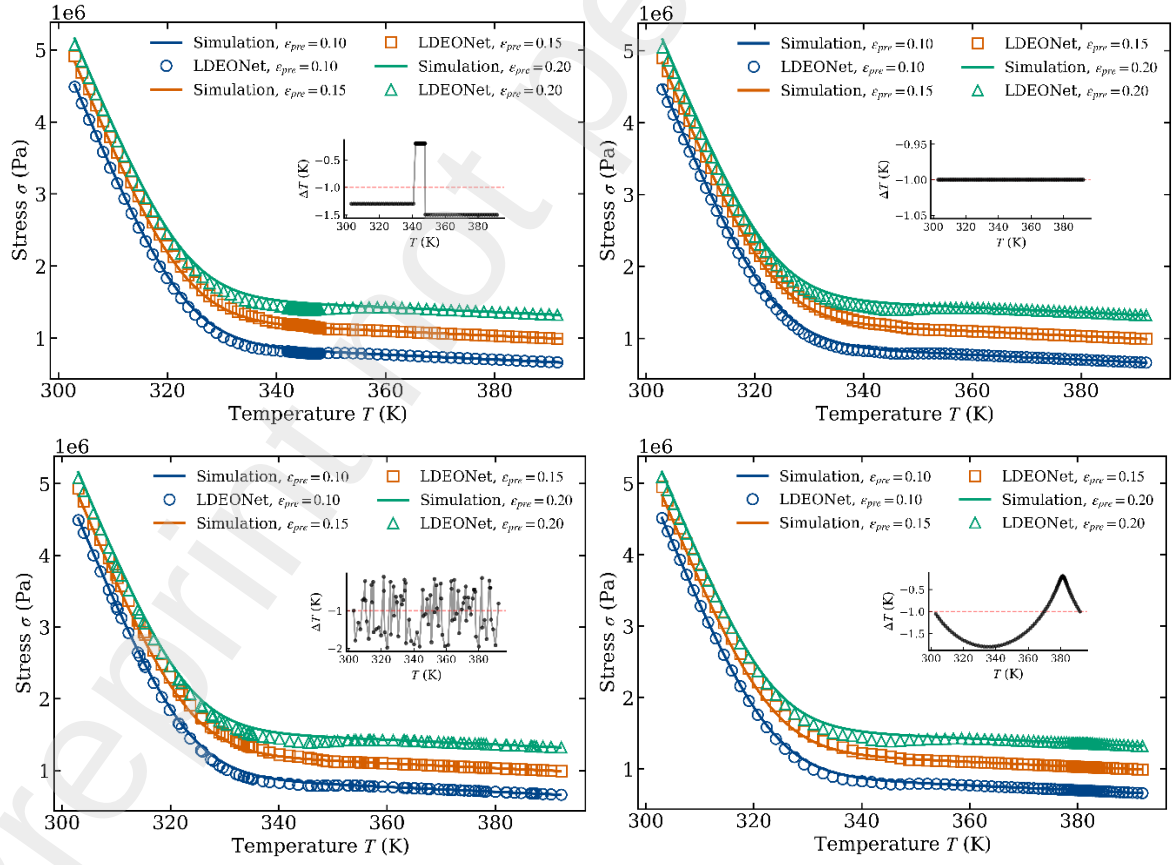


Figure 9: Discretization generalization performance: (a) Constant baseline; (b) Random noise; (c) Sinusoidal fluctuation; (d) Stepwise variation

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

The discretization generalization capability of LDEONet also implies another important practical value: it can be seamlessly integrated with arbitrary numerical integration schemes (e.g., Euler method, Runge-Kutta methods), and improvements in integration accuracy can be directly achieved by refining the step size without necessitating model retraining. This characteristic offers distinct advantages for embedding LDEONet within the adaptive solvers of finite element software.

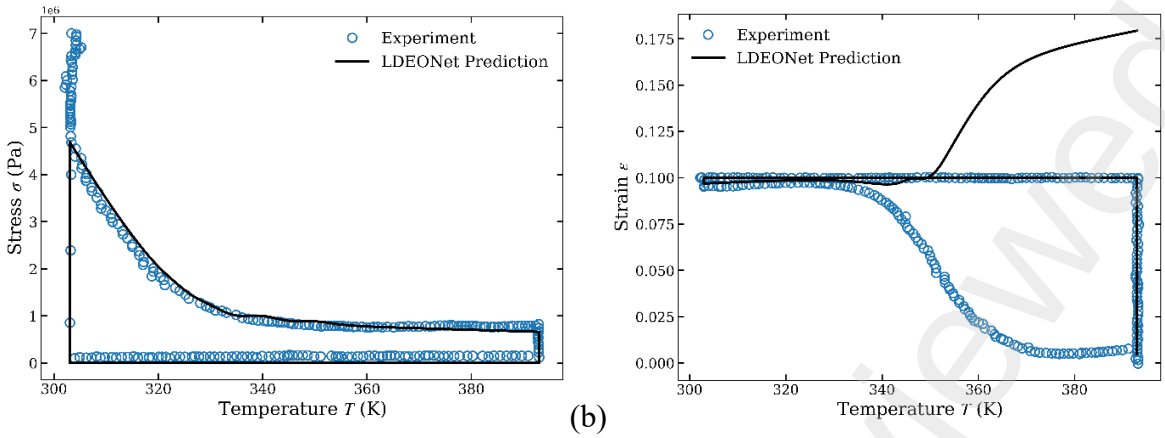
4.3 Mode Generalization Performance

Under typical service conditions, SMPs generally undergo complete coupled thermomechanical cycles encompassing high-temperature loading, cooling and fixing, isothermal unloading, and heating recovery. Having verified that LDEONet-A4 possesses path generalization and discretization generalization capabilities, this subsection further investigates the predictive capacity of LDEONet-A4 for unknown physical evolution stages after being trained exclusively on a single evolution stage. This capacity is herein defined as mode generalization, and assessing it is crucial for the engineering application of neural network constitutive models under complex, non-stationary operating conditions.

To this end, a mode generalization performance experiment is designed, utilizing the full-cycle evolution process of SMPs to examine the mode generalization performance of LDEONet-A4. In the specific implementation, a pre-strain of 0.15 is initially applied and maintained at an initial temperature of $T_0 = 393\ \mathrm{K}$. Subsequently, the system is cooled at a constant strain with a step size of $\Delta T = -1\ \mathrm{K}$ to $T_1 = 303\ \mathrm{K}$. Thereafter, complete unloading is performed until the macroscopic stress reaches $\sigma = 0\ \mathrm{MPa}$. Finally, continuous heating is applied under unconstrained conditions to drive free strain recovery.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734



(a) (b)

Figure 10: Mode generalization performance: (a) Temperature--strain curve; (b) Temperature--stress curve

As illustrated in **Figure 10**, the performance test results indicate that although LDEONet achieves high predictive accuracy during the cooling and fixing stage, faithfully reconstructing the evolution trajectories of strain with respect to temperature and stress, it exhibits pronounced generalization failure during the heating recovery stage. The strain predicted by the model does not decay to zero as the temperature increases. Instead, it displays an anomalous sustained increasing trend, and its evolution trajectory fundamentally deviates from the genuine physical process.

The physical origin of this failure lies in the thermomechanical path asymmetry inherent in the internal variable evolution mechanism of SMPs. Specifically, during the cooling and fixing stage, the material is subjected to a constrained loading state, and the accumulation of stored strain is primarily driven by the partial freezing of elastic strain. The corresponding physical evolution operator can be approximately expressed as:

$$D_{\text{cooling}} = \boldsymbol{S}_r : \boldsymbol{\sigma} \tag{22}$$

In this mode, the evolution direction exhibits a strong correlation with the elastic strain $\boldsymbol{\varepsilon}_e$. Conversely, during the heating recovery stage, the macroscopic stress is completely removed, and the system resides under free boundary conditions. Under these circumstances, the release of stored strain is no longer driven by external stress but is instead governed by the conformational entropy increase mechanism accompanying the transition from the glassy phase to the rubbery phase. The dominant physical operator undergoes a modal switch and is typically characterized as:

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

$$D_{\text{heating}} = \frac{\boldsymbol{\varepsilon}_s}{\phi_g} \tag{23}$$

This failure phenomenon reveals the limitations of the basic LDEONet architecture when confronting switches in material evolution modes. Relying solely on the laws derived from a single evolution mode renders it difficult to accommodate multi-mode history-dependent materials such as SMPs. To overcome the generalization bottleneck imposed by mode asymmetry, the LDEONet architecture may require the explicit introduction of a perception mechanism capable of characterizing microstructural mechanism switches, or the incorporation of cross-mode thermodynamic consistency penalties to guide the network toward constructing a global physical operator applicable to the full-cycle path.

4.4 Model Comparison

To systematically evaluate the performance and characteristics of the LDEONet architecture, this subsection first compares it with four representative model classes from the perspective of history-dependent material modeling at both theoretical and architectural levels. The comparison encompasses purely data-driven sequential models (RNNs, LSTM) [49], deep learning models incorporating physical constraints (PINNs, TANNs) [50], Internal State Variable Neural Ordinary Differential Equations (ISV-NODE) based on continuous-time integration [40], and the History-Aware Neural Operator (HANO) oriented toward operator learning [48].

From the architectural benchmark of model scale, the parameter count of LSTM typically resides on the order of 10^5, HANO on the order of 10^4, ISV-NODE on the order of 10^3, and PINNs on the order of 10^2. In contrast, LDEONet, through local operator decoupling, compresses the parameter count to the order of 10^1, demonstrating a pronounced lightweight advantage. Regarding path generalization capability, LSTM, HANO, and ISV-NODE all rely on complete or partial historical sequence features for training or prediction. Although conventional PINN is equation-driven, its solution process often implicitly couples the global evolution domain with specific initial conditions. These characteristics may limit the ability of existing models to directly initiate predictions from arbitrary intermediate states without

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

additional mechanisms such as state reconstruction or retraining. By comparison, LDEONet is dedicated to approximating the local evolution law of internal variables and is thus inherently endowed with path generalization capability. Concerning discretization generalization capability, PINNs and ISV-NODE theoretically possess this capacity by virtue of their continuous-time formulation. HANO exhibits partial adaptability to sampling windows. LSTM is constrained by the strong coupling between recurrent weights and fixed sampling intervals. LDEONet, owing to the explicit decoupling of evolution direction and evolution intensity, restricts the network to mapping solely the evolution direction per unit phase transition, while variations in the external discretization step size act exclusively upon the numerical integrator. Consequently, discretization generalization capability is realized at the architectural level. With respect to physical interpretability and training cost, PINN directly embeds partial differential equations, affording the strongest interpretability but often accompanied by convergence difficulties arising from non-convex optimization. LSTM and HANO tend toward high-dimensional black-box mappings with weaker interpretability. The design of LDEONet corresponds strictly to the structural decomposition of the internal variable evolution equation, offering a clear physical interpretation. Moreover, owing to its singular learning objective, the training process of LDEONet is stable and converges rapidly.

Focusing further on the cooling and fixing process of SMPs, which constitutes the central theme of this study, the quantitative performance metrics of three representative models in practical applications are compared in **Table 2**. For the constitutive prediction of SMPs, existing models exhibit distinctly different modeling philosophies in terms of accuracy, data requirements, and generalization flexibility. LSTM, leveraging its deep sequential memory mechanism, achieves a comparatively low mean squared error (on the order of 10^{-5}). However, this representational capacity is predicated on a massive parameter scale and the training cost associated with extensive long-sequence samples (3680 curves), while sacrificing generalization flexibility for complex loading scenarios. The PINN framework achieves a mean squared error on the order of 10^{-4} with an extremely small sample size (2.7 curves), demonstrating the powerful constraining capability of physical priors. Nonetheless, constrained by its characteristic end-to-end learning of the temperature--stress relationship, it

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

still lacks path generalization capability.

In comparison, LDEONet, under test conditions utilizing only 21 curves and with a substantially reduced parameter count, stably converges to an accuracy level comparable to that of PINN at a relatively low iteration cost (~10^4 iterations). Owing to the advantages conferred by its theoretical architecture, LDEONet stands as the sole model among the three capable of initiating continuous predictions from arbitrary state points with arbitrary time step sizes in practical engineering simulations without requiring retraining.

5. Conclusion

This study addresses the common challenge wherein existing neural network constitutive models for history-dependent materials exhibit excessive reliance on global loading sequences, and re-examines the physical essence of internal variable evolution. Through derivation and extension, this work elucidates that the internal variable evolution law of CHDMs, exemplified by SMPs, possesses the Markovian property. Building upon this theoretical foundation, LDEONet is innovatively proposed, and the generalization performance of the LDEONet-A4 model is rigorously tested and validated through comprehensive generalization experiments. The principal conclusions and contributions of this study are summarized as follows:

(a) Revelation of the Markovian property of the evolution process.

This study theoretically demonstrates that, for CHDMs, the macroscopic history-dependent memory effects are completely encoded and stored within the current finite-dimensional internal state variables. The thermomechanical evolution of the material inherently obeys a local internal variable evolution law, rather than exhibiting explicit dependence on long-range loading paths. Consequently, existing neural network models that rely on historical sequential features—such as RNNs, LSTMs, and HANO—exhibit structural redundancy when applied to such local evolution problems.

(b) Proposal of the LDEONet architecture and theoretical advances in sample efficiency.

Based on the internal variable evolution law, this study proposes the Local Decoupling

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

Evolution Operator Network. This framework explicitly decouples the complex material evolution law into a local evolution direction and an evolution intensity. The investigation confirms that LDEONet achieves high-accuracy convergence with only a minimal number of samples while compressing the model parameter count to the order of 10^1, thereby demonstrating progress in both sample efficiency and lightweight modeling.

(c) Validation of model generalization capability and exploration of model limitations.

Through the design of multi-dimensional generalization performance tests, this study verifies that LDEONet is capable of initiating continuous predictions from arbitrary intermediate states with arbitrary discrete time steps, simultaneously attaining both path generalization and discretization generalization capabilities. Furthermore, this study thoroughly examines the generalization failure encountered by the model when the evolution mechanism undergoes a modal switch, delineates the boundaries of the generalization capability of LDEONet, and articulates its key limitations.

In summary, the research presented herein not only furnishes a constitutive prediction tool for history-dependent materials that reconciles lightweight architecture with high generalization capability, but more importantly, establishes a novel paradigm for neural network constitutive modeling of history-dependent materials—namely, leveraging the underlying local evolution mechanisms of the material to construct more generalizable models. Although LDEONet exhibits generalization failure when confronting cross-modal transitions, this framework substantiates the considerable research value of aligning model architecture with fundamental physical laws. Future investigations will concentrate on extending LDEONet to finite deformation frameworks in three-dimensional settings and on its deep integration with finite element solution frameworks, thereby advancing the application of neural network constitutive models to high-performance nonlinear numerical simulations.

References

[1] S. K. Chandaka, A. Das, P. Laskar, Emergence of shape memory polymers as a new material for diverse applications, RSC Advances 15 (2025) 31210 – 31229.

[2] Y. Teng, Z. Shi, S. Hou, Q. Lv, J. Shi, H. Fan, S. Ramakrishna, Intelligent/Smart materials for wearables used in fire emergency medical services, International Journal

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

of Smart and Nano Materials 16 (2025) 726 – 741. doi:10.1080/19475411.2025.2559388, review Article.

[3] A. A. Fetisova, M. A. Surmeneva, R. A. Surmenev, Advanced Design Concepts for Shape-Memory Polymers in Biomedical Applications and Soft Robotics, Polymers 18 (2026) 214.

[4] L. Flórido Martins, C. A. Spiegel, E. Blasco, 4D microprinting of dual shape and triple shape memory polymers, Cell Reports Physical Science 7 (2026) 103137.

[5] S. Dutta, S. Havanagi, P. Mishra, D. Sharma, B. N. Singh, From inert to intelligent: smart biomaterials reshaping the landscape of bioengineering, International Journal of Smart and Nano Materials 16 (2025) 845 – 877. doi:10.1080/19475411.2025.2570535, review Article.

[6] X. Zhao, C. Chang, W. Zhu, A. Ablali, M. Li, F. Xie, Y. Li, Energy harvesting dynamic analysis of a piezoelectric beam under axially moving and spinning motions based on the Green’s functions, International Journal of Smart and Nano Materials 16 (2025) 878 – 907. doi:10.1080/19475411.2025.2574076, research Article.

[7] J. Deng, H. Lu, Y. Fu, A Physically-Based Thermo-Viscoelastic Constitutive Model of Phantom and Entangled Networks in Amorphous Shape Memory Polymers, SSRN Preprint, 2024. URL: https://www.ssrn.com/abstract=5067758. doi:10.2139/ssrn.5067758.

[8] C. B. Cooper, S. Nikzad, H. Yan, Y. Ochiai, J.-C. Lai, Z. Yu, G. Chen, J. Kang, Z. Bao, High Energy Density Shape Memory Polymers Using Strain-Induced Supramolecular Nanostructures, ACS Central Science 7 (2021) 1657 – 1667. URL: https://pubs.acs.org/doi/10.1021/acscentsci.1c00829. doi:10.1021/acscentsci.1c00829.

[9] Z. Xu, S. Meng, D.-W. Wei, R.-Y. Bao, Y. Wang, K. Ke, W. Yang, Hierarchical network relaxation of a dynamic cross-linked polyolefin elastomer for advanced reversible shape memory effect, Nanoscale 15 (2023) 5458 – 5468. URL: https://xlink.rsc.org/?DOI=D2NR06902D. doi:10.1039/D2NR06902D.

[10] Y.-C. Chen, D. C. Lagoudas, A constitutive theory for shape memory polymers. Part I, Journal of the Mechanics and Physics of Solids 56(2008) 1752 – 1765. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022509607002323. doi:10.1016/j.jmps.2007.12.005.

[11] H. Tobushi, K. Okumura, S. Hayashi, N. Ito, Thermomechanical constitutive model of shape memory polymer, Mechanics of Materials 33(2001) 545 – 554. URL:https://www.sciencedirect.com/science/article/pii/S0167663601000758. doi:https://doi.org/10.1016/S0167-6636(01)00075-8.

[12] W. Zhao, L. Liu, J. Leng, Y. Liu, Thermo-mechanical behavior prediction of shape memory polymer based on the multiplicative decomposition of the deformation gradient, Mechanics of Materials 143 (2020) 103263. URL: https://linkinghub.elsevier.com/retrieve/pii/S0167663619304260. doi:10.1016/j.mechmat.2019.103263.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

[13] H. Park, P. Harrison, Z. Guo, M.-G. Lee, W.-R. Yu, Three-dimensional constitutive model for shape memory polymers using multiplicative decomposition of the deformation gradient and shape memory strains, Mechanics of Materials 93 (2016) 43 – 62. URL: https://linkinghub.elsevier.com/retrieve/pii/S0167663615002239. doi:10.1016/j.mechmat.2015.10.014.

[14] B. L. Volk, D. C. Lagoudas, D. J. Maitland, Characterizing and modeling the free recovery and constrained recovery behavior of a polyurethane shape memory polymer, Smart Materials and Structures 20 (2011) 094004. URL: https://iopscience.iop.org/article/10.1088/0964-1726/20/9/094004. doi:10.1088/0964-1726/20/9/094004.

[15] B. L. Volk, D. C. Lagoudas, Y.-C. Chen, Analysis of the finite deformation response of shape memory polymers: II. 1D calibration and numerical implementation of a finite deformation, thermoelastic model, Smart Materials and Structures 19 (2010) 075006. URL: https://iopscience.iop.org/article/10.1088/0964-1726/19/7/075006. doi:10.1088/0964-1726/19/7/075006.

[16] F. Li, Constitutive model and space application verification of epoxy-based shape memory polymer composite, Ph.D. thesis, Harbin Institute of Technology, Harbin, 2020. doi:10.27061/d.cnki.ghgdu.2020.004918.

[17] J. N. Fuhg, G. Anantha Padmanabha, N. Bouklas, B. Bahmani, W. Sun, N. N. Vlassis, M. Flaschel, P. Carrara, L. De Lorenzis, A Review on Data-Driven Constitutive Laws for Solids, Archives of Computational Methods in Engineering 32 (2025) 1841 – 1883. URL: https://link.springer.com/10.1007/s11831-024-10196-2. doi:10.1007/s11831-024-10196-2.

[18] D. W. Abueidda, S. Koric, N. A. Sobh, H. Sehitoglu, Deep learning for plasticity and thermo-viscoplasticity, International Journal of Plasticity 136 (2021) 102852. URL: https://linkinghub.elsevier.com/retrieve/pii/S0749641920302096. doi:10.1016/j.ijplas.2020.102852.

[19] F. Feyel, A multilevel finite element method (FE2) to describe the response of highly non-linear structures using generalized continua, Computer Methods in Applied Mechanics and Engineering 192 (2003) 3233 – 3244. URL: https://linkinghub.elsevier.com/retrieve/pii/S0045782503003487. doi:10.1016/S0045-7825(03)00348-7.

[20] N. A. Collins-Craft, I. Stefanou, J. Sulem, I. Einav, A Cosserat Breakage Mechanics model for brittle granular media, Journal of the Mechanics and Physics of Solids 141 (2020) 103975. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022509620302106. doi:10.1016/j.jmps.2020.103975.

[21] F. Masi, I. Stefanou, P. Vannucci, A study on the effects of an explosion in the Pantheon of Rome, Engineering Structures 164 (2018) 259 – 273. URL: https://linkinghub.elsevier.com/retrieve/pii/S0141029617309501. doi:10.1016/j.engstruct.2018.02.082.

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

[22] F. Masi, I. Stefanou, V. Maffi-Berthier, P. Vannucci, A Discrete Element Method based-approach for arched masonry structures under blast loads, Engineering Structures 216 (2020) 110721. URL: https://linkinghub.elsevier.com/retrieve/pii/S0141029619328585. doi:10.1016/j.engstruct.2020.110721.
[23] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Computation 9 (1997) 1735 – 1780. doi:10.1162/neco.1997.9.8.1735.
[24] M. Schuster, K. Paliwal, Bidirectional recurrent neural networks, IEEE Transactions on Signal Processing 45 (1997) 2673 – 2681. URL: http://ieeexplore.ieee.org/document/650093/. doi:10.1109/78.650093.
[25] A. Graves, J. Schmidhuber, Framewise phoneme classification with bidirectional LSTM and other neural network architectures, Neural Networks 18 (2005) 602 – 610. URL: https://linkinghub.elsevier.com/retrieve/pii/S0893608005001206. doi:10.1016/j.neunet.2005.06.042.
[26] A. Graves, A.-r. Mohamed, G. Hinton, Speech recognition with deep recurrent neural networks, in: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE, Vancouver, BC, Canada, 2013, pp. 6645 – 6649. URL: http://ieeexplore.ieee.org/document/6638947/. doi:10.1109/ICASSP.2013.6638947.
[27] A. Graves, N. Jaitly, A.-r. Mohamed, Hybrid speech recognition with Deep Bidirectional LSTM, in: 2013 IEEE Workshop on Automatic Speech Recognition and Understanding, IEEE, Olomouc, Czech Republic, 2013, pp. 273 – 278. URL: http://ieeexplore.ieee.org/document/6707742/. doi:10.1109/ASRU.2013.6707742.
[28] Z. Huang, W. Xu, K. Yu, Bidirectional LSTM-CRF Models for Sequence Tagging, 2015. URL: https://arxiv.org/abs/1508.01991. doi:10.48550/arXiv.1508.01991. arXiv:1508.01991, arXiv preprint.
[29] M. Oeser, S. Freitag, Modeling of materials with fading memory using neural networks, International Journal for Numerical Methods in Engineering 78 (2009) 843 – 862. URL: https://onlinelibrary.wiley.com/doi/10.1002/nme.2518. doi:10.1002/nme.2518.
[30] W. Graf, S. Freitag, M. Kaliske, J.-U. Sickert, Recurrent Neural Networks for Uncertain Time-Dependent Structural Behavior: Recurrent neural networks for uncertain time-dependent structural behavior, Computer-Aided Civil and Infrastructure Engineering 25 (2010) 322 – 323. URL: https://onlinelibrary.wiley.com/doi/10.1111/j.1467-8667.2009.00645.x. doi:10.1111/j.1467-8667.2009.00645.x.
[31] W. Graf, S. Freitag, J. Sickert, M. Kaliske, Structural Analysis with Fuzzy Data and Neural Network Based Material Description, Computer-Aided Civil and Infrastructure Engineering 27 (2012) 640 – 654. URL: https://onlinelibrary.wiley.com/doi/10.1111/j.1467-8667.2012.00779.x. doi:10.1111/j.1467-8667.2012.00779.x.
[32] S. Freitag, W. Graf, M. Kaliske, Recurrent neural networks for fuzzy data, Integrated Computer-Aided Engineering 18 (2011) 265 – 280. URL:

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

https://journals.sagepub.com/doi/full/10.3233/ICA-2011-0373. doi:10.3233/ICA-2011-0373.
[33] J. Dornheim, L. Morand, H. J. Nallani, D. Helm, Neural Networks for Constitutive Modeling: From Universal Function Approximators to Advanced Models and the Integration of Physics, Archives of Computational Methods in Engineering 31 (2024) 1097 – 1127. URL: https://link.springer.com/10.1007/s11831-023-10009-y. doi:10.1007/s11831-023-10009-y.
[34] P. Zhang, Z. Yin, Y. Jin, G. Ye, An AI-based model for describing cyclic characteristics of granular materials, International Journal forNumerical and Analytical Methods in Geomechanics 44 (2020) 1315 – 1335. URL: https://onlinelibrary.wiley.com/doi/10.1002/nag.3063. doi:10.1002/nag.3063.
[35] M. B. Gorji, M. Mozaffar, J. N. Heidenreich, J. Cao, D. Mohr, On the potential of recurrent neural networks for modeling path dependent plasticity, Journal of the Mechanics and Physics of Solids 143 (2020) 103972. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022509620302076. doi:10.1016/j.jmps.2020.103972.
[36] S. Anumasa, G. Gunapati, P. K. Srijith, Continuous Depth Recurrent Neural Differential Equations, in: D. Koutra, C. Plant, M. Gomez Rodriguez, E. Baralis, F. Bonchi (Eds.), Machine Learning and Knowledge Discovery in Databases: Research Track, Springer Nature Switzerland, Cham, 2023, pp. 223 – 238.
[37] P. Sun, J. Wu, M. Zhang, P. Devos, D. Botteldooren, Delayed memory unit: Modeling temporal dependency through delay gate, IEEE Transactions on Neural Networks and Learning Systems 36 (2025) 10808 – 10818.
[38] S. Kapturowski, G. Ostrovski, W. Dabney, J. Quan, R. Munos, Recurrent experience replay in distributed reinforcement learning, in: International Conference on Learning Representations, 2019. URL: https://openreview.net/forum?id=r1lyTjAqYX.
[39] V. Tac, F. Sahli Costabal, A. B. Tepole, Data-driven tissue mechanics with polyconvex neural ordinary differential equations, Computer Methods in Applied Mechanics and Engineering 398 (2022) 115248. URL: https://linkinghub.elsevier.com/retrieve/pii/S0045782522003838. doi:10.1016/j.cma.2022.115248.
[40] R. E. Jones, J. N. Fuhg, An attention-based neural ordinary differential equation framework for modeling inelastic processes, 2025. URL: https://arxiv.org/abs/2502.10633. arXiv:2502.10633.
[41] Y. He, S. J. Semnani, Incremental Neural Controlled Differential Equations for modeling of path-dependent material behavior, Computer Methods in Applied Mechanics and Engineering 422 (2024) 116789. URL: https://linkinghub.elsevier.com/retrieve/pii/S0045782524000458. doi:10.1016/j.cma.2024.116789.
[42] F. Masi, I. Stefanou, P. Vannucci, V. Maffi-Berthier, Thermodynamics-based Artificial Neural Networks for constitutive modeling, Journal of the Mechanics and Physics of

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

Solids 147 (2021) 104277. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022509620304841. doi:10.1016/j.jmps.2020.104277.

[43] F. Masi, I. Stefanou, Multiscale modeling of inelastic materials with Thermodynamics-based Artificial Neural Networks (TANN), Computer Methods in Applied Mechanics and Engineering 398 (2022) 115190. URL: https://linkinghub.elsevier.com/retrieve/pii/S0045782522003450. doi:10.1016/j.cma.2022.115190.

[44] D. K. Klein, M. Fernández, R. J. Martin, P. Neff, O. Weeger, Polyconvex anisotropic hyperelasticity with neural networks, Journal of the Mechanics and Physics of Solids 159 (2022) 104703. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022509621003215. doi:10.1016/j.jmps.2021.104703.

[45] P. Thakolkaran, A. Joshi, Y. Zheng, M. Flaschel, L. De Lorenzis, S. Kumar, NN-EUCLID: Deep-learning hyperelasticity without stress data, Journal of the Mechanics and Physics of Solids 169 (2022) 105076. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022509622002538. doi:10.1016/j.jmps.2022.105076.

[46] L. Lu, P. Jin, G. Pang, Z. Zhang, G. E. Karniadakis, Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators, Nature Machine Intelligence 3 (2021) 218 – 229. URL: https://doi.org/10.1038/s42256-021-00302-5. doi:10.1038/s42256-021-00302-5.

[47] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, A. Anandkumar, Fourier neural operator for parametric partial differential equations, 2021. URL: https://arxiv.org/abs/2010.08895. arXiv:2010.08895.

[48] B. Guo, Z. Lin, Q. He, History-aware neural operator: Robust data-driven constitutive modeling of path-dependent materials, Computer Methods in Applied Mechanics and Engineering 447 (2025) 118358. URL: https://linkinghub.elsevier.com/retrieve/pii/S0045782525006309. doi:10.1016/j.cma.2025.118358.

[49] J. Deng, H. Lu, Deep learning of long short-term memory neural networks in shape memory polymers towards shape memory behaviors, Polymer 346 (2026) 129651. URL: https://linkinghub.elsevier.com/retrieve/pii/S0032386126000984. doi:10.1016/j.polymer.2026.129651.

[50] C. Yan, X. Feng, P. Mensah, G. Li, Toward a general physics-informed neural network for amorphous shape memory polymer modelling, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 481 (2025) 20240702. URL: https://doi.org/10.1098/rspa.2024.0702. doi:10.1098/rspa.2024.0702.

[51] Y. Liu, K. Gall, M. L. Dunn, A. R. Greenberg, J. Diani, Thermomechanics of shape memory polymers: Uniaxial experiments and constitutive modeling, International Journal of Plasticity 22 (2006) 279 – 313. URL:

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734

https://linkinghub.elsevier.com/retrieve/pii/S0749641905000604.
doi:10.1016/j.ijplas.2005.03.004.

Table 1: Training results under different hidden layer configurations

Model	Configuration	MSE	R^2
LDEONet-A1	1 hidden layer, 5 neurons	1.5747×10^{-2}	0.9833
LDEONet-A2	1 hidden layer, 10 neurons	1.0708×10^{-3}	0.9993
LDEONet-A3	1 hidden layer, 20 neurons	1.1158×10^{-3}	0.9991
LDEONet-A4	1 hidden layer, 30 neurons	7.3167×10^{-4}	0.9997

Table 2: Comprehensive performance comparison of neural network constitutive models for shape memory polymers

Metric	LSTM	PINN	LDEONet (This Work)
Parameter scale	$\sim10^{5}$	$\sim10^{2}$	$\sim10^{1}$
Training sample size	3680 curves	2.7 curves	21 curves
Convergence accuracy (MSE)	$\sim10^{-5}$	$\sim10^{-4}$	$\sim10^{-4}$
Optimization iterations	~500	$\sim10^{6}$	$\sim10^{4}$
Path generalization	Not supported	Not supported	Supported
Discretization generalization	Not supported	Not supported	Supported

Preprint not peer reviewed

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=6996734