International Journal of Solids and Structures 334 (2026) 113989

Contents lists available at ScienceDirect

International Journal of Solids and Structures

journal homepage: www.elsevier.com/locate/ijsolstr

A spatiotemporal framework for strain gradient viscoelasticity: Thermodynamic derivation and numerical implementation

Zhongya Lin [a,1], Yanfei Wang [b,*], Yueguang Wei [a,*]

[a] Department of Mechanics and Engineering Science, College of Engineering, Peking University, Beijing 100871, China
[b] State Key Laboratory for Strength and Vibration of Mechanical Structures, Department of Engineering Mechanics, Xi'an Jiaotong University, Xi'an 710049, China

ARTICLE INFO

Keywords:
Strain gradient viscoelasticity
Spatiotemporal behavior
Mixed finite element
Physics-informed neural networks

ABSTRACT

The mechanical response of many advanced materials is governed by a complex interplay between microstructural size effects and time-dependent dissipative processes. Classical continuum theories are insufficient for these systems as they lack intrinsic length and time scales. This paper introduces strain gradient viscoelasticity (SGV) within the thermoviscoelastic framework, which is a higher-order continuum theory that provides a unified spatiotemporal framework linking spatial gradients to temporal evolution. We develop and validate two distinct numerical solution schemes: a robust mixed-field finite element method and a physics-informed neural networks framework, which naturally handles higher-order derivatives via automatic differentiation and manages the time integration through a recurrence time-stepping formula. Using the beam bending problem, we demonstrate that both numerical models show excellent agreement with analytical solutions. Then, by constructing spatiotemporal phase portraits, we visualize the coupled nature of the theory, demonstrating how microstructural features and viscosity fundamentally modulate the mechanical property. Furthermore, we demonstrate the potential applicability of SGV by showing its ability to characterize the multiscale viscoelasticity of biological tissues, resolving distinct relaxation times associated with both cellular and microstructural components. This study provides a theoretical and computational foundation for modeling advanced materials where scale and time evolution are inextricably linked.

1. Introduction

The quest to unravel the intricate mechanical behaviors of advanced materials, from nanostructured crystalline materials to biological tissues, remains a central theme in various engineering and scientific domains. In engineered nanomaterials, such as metallic glasses and thermal barrier coatings (Liu et al., 2018; An et al., 2020; Zhu et al., 2018), microstructural heterogeneities interact with time-dependent processes like creep or damping. Similarly, in biological systems, the embryo morphogenesis and tissue development are governed by the interplay of elasticity and viscosity across multiple scales (Hadzipasic et al., 2024; Mongera et al., 2023). These materials often operate from the microscale to macroscale. The continuum theory is needed which can account for intrinsic length scales and temporal dependencies that govern phenomena like size-dependent stiffness, viscous relaxation, and dynamic evolution process.

Classical continuum theories, while remarkably successful for many macroscopic applications, often rely on scale-invariance assumptions, overlooking the gradient effects arising from microstructural features such as interfaces, surface energy, and molecular architecture. As the representative microstructure of these materials decreases into the micro- and nano-scale, the limitations of classical continuum models become more pronounced. To address this deficiency, strain gradient continuum theories have emerged as a powerful framework. By incorporating higher-order gradients of the displacement or strain tensor into the constitutive relations and introducing intrinsic material length scale parameters, these theories can quantify the influence of microstructure on the macroscopic mechanical response (Askes and Aifantis, 2011; Mindlin, 1965; Mindlin and Eshel, 1968; Wei and Hutchinson, 1997). The theoretical foundation of gradient theories can be traced back to the seminal works of Mindlin (Mindlin, 1965, 1964), who introduced the general theory of elasticity with microstructure. Then, various formulations were developed to resolve the limitations of classical continuum mechanics at microscales. These formulations range from theories based

* Corresponding authors.
E-mail addresses: yfwang@xjtu.edu.cn (Y. Wang), weiyg@pku.edu.cn (Y. Wei).
[1] Present address: Aix-Marseille Université & CNRS, IBDM - UMR7288, 13009 Marseille, France.

https://doi.org/10.1016/j.ijsolstr.2026.113989
Received 11 November 2025; Received in revised form 20 March 2026; Accepted 31 March 2026
Available online 1 April 2026

on the second gradient of displacement to those incorporating independent microrotations (Chen and Wang, 2002; Reddy, 2007; Yu et al., 2021). The gradient of the strain tensor is a robust kinematic variable for capturing size effects (Askes et al., 2008; Ding et al., 2023; Polyzos and Fotiadis, 2012). The higher-order terms in gradient theories are essential not only for modeling size-dependent stiffness but also for the regularization of singularities in defect mechanics (Askes and Aifantis, 2011; Lazar et al., 2006). Moreover, recent thermodynamic analysis has demonstrated that the strain gradient provides a physically consistent work-conjugate for higher-order stresses (Polizzotto, 2016). Strain gradient theories have also been developed to derive second-order gradient constitutive models for architectural materials, establishing a link between discrete microstructural mechanics and equivalent continuum responses (Rahali et al., 2021; Rahali et al., 2017). Concurrently, many advanced materials, such as biomaterials and metals with nanostructures, exhibit viscous behavior (Chen et al., 2022; Ebrahimi and Barati, 2017; Lin et al., 2025a). The viscosity of polymers or hydrogels can come from network entanglement (Gusev and Bernhard, 2024) and chain interaction (Huang et al., 2022). Grain boundary slips as well as thermally-activated dislocation movement can result in viscous behavior in metallic nanocrystalline materials. To accurately characterize these material behaviors, the constitutive theory should explicitly account for the coupling between microstructural spatial gradients and temporal evolution.

Integrating spatial and temporal effects can be achieved by formulating the strain gradient theory within the framework of irreversible thermodynamics. By introducing the strain gradient into the free energy density and defining appropriate dissipation potentials, the resulting constitutive law naturally accounts for both higher-order spatial stresses and viscous effects. Recent advancements have extended strain gradient model to fracture mechanics, where full-field solutions for cracks reveal time-evolving stress concentrations mitigated by length scales (Ding et al., 2023), and to polymer networks, capturing entanglement effects through gradient-dependent relaxation moduli (Jiang et al., 2023). In this study, we propose the strain gradient viscoelasticity (SGV) using the thermoviscoelastic framework, which respects the first and second laws of thermodynamics to ensure energy conservation and positive entropy production. The proposed theory results in a system of higher-order partial differential equations where the spatial gradients are coupled with the temporal evolution, thereby linking microstructural deformation to macroscopic response. This would be useful to characterize the evolution of nonstructural crystalline materials, the morphological development of active biological tissues, and other complex, time-evolving processes where scale and history are intertwined.

Furthermore, the mathematical formulation of SGV typically results in governing partial differential equations that involve both higher-order spatial derivatives and temporal derivatives or integrals. Solving these complex equations analytically may be intractable, except for simple structures. Therefore, robust numerical methods are indispensable. The finite element method stands as the workhorse for computational solid mechanics, widely adapted for classical viscoelastic problems, although its extension to strain gradient theories presents non-trivial challenges. Standard finite element formulations often struggle with the C^1 continuity requirement imposed by second or higher gradients of displacement, necessitating the development of specialized complex finite elements (Wei, 2006), or adding node freedoms to decrease the continuity requirement (Zybell et al., 2012). Implementing these within a time-dependent viscoelastic framework further compounds the complexity. In recent years, physics-informed neural networks (PINNs) have emerged as a new computational method (Bai et al., 2025; Karniadakis et al., 2021), leveraging the expressive power of deep learning to solve differential equations. PINNs embed the governing physical laws, boundary conditions, and initial conditions directly into the loss function of a neural network (Grossmann et al., 2024; Wang et al., 2024). This approach offers potential advantages for problems involving higher-order derivatives, which can be handled more naturally through automatic differentiation inherent in neural network training.

This study advances the field by presenting a comprehensive computational investigation of SGV, with a dual focus on elucidating its spatiotemporal capabilities and comparing state-of-the-art numerical solution strategies. We develop a finite element model for SGV, addressing continuity and stability issues through mixed elements and efficient time integration. Furthermore, we pioneer a PINNs framework for SGV, demonstrating its prowess in capturing the higher-order boundary conditions and resolving time integration by recurrence formula. Then, we provide a comparison of these two numerical results and the analytical solution using the beam bending problem. Through this work, we present both fundamental insights and practical computational guidance for tackling mechanical problems where the interplay of microstructural scale and time-dependency is crucial. This paper is structured as follows: Section 2 develops the strain gradient viscoelasticity with the first and second law of thermodynamics. Section 3 details the formulation and implementation using the mixed finite element method and PINNs. Section 4 presents results and a comparative analysis of the two numerical methods. Section 5 introduces the application of SGV to a biophysical problem, which can reveal the multiscale viscoelasticity mechanisms. Finally, Section 6 provides concluding remarks and discusses future research directions.

2. Thermoviscoelastic framework for strain gradient viscoelasticity

Classical linear viscoelasticity describes the time-dependent behavior of polymers, glasses, and many engineering materials through stress–strain relations that involve hereditary integrals or differential rheological models. However, when the deformation field varies strongly over small spatial scales, classical continuum theories become inadequate because they do not incorporate intrinsic material length scales (Huang et al., 2020; Wang and Wei, 2024). Strain gradient viscoelasticity can provide a natural extension of classical viscoelasticity by including the dependence of the free energy and dissipation on the gradients of strain in addition to the strain itself. This extension introduces internal length parameters that regularize localization and endow the model with objective size effects. A rigorous way to develop a consistent strain-gradient viscoelastic theory is to derive it within the thermoviscoelastic framework, starting from the local balances of mass, momentum, and energy together with the Clausius–Duhem inequality. This ensures that both the classical stress and the higher-order stress obey constitutive laws compatible with thermodynamics.

Following the thermoviscoelastic formulation in classical continuum mechanics, the local form of the energy balance for infinitesimal deformation can be written as (Christensen, 1982)

$$\rho\dot{e} = \dot{W} + \rho r - Q_{i,i}, \tag{1}$$

where ρ is mass density, e is internal energy per unit mass, r is heat supply per unit mass, Q_i is heat flux per unit area per unit time, comma in the subscript denotes partial derivative, and W is the deformation work per unit volume, which can be written as

$$\dot{W} = \sigma_{ij}\dot{\varepsilon}_{ij} + \tau_{ijk}\dot{\varepsilon}_{ijk}, \tag{2}$$

where is σ_{ij} the Cauchy stress, $\varepsilon_{ij} = 1/2(u_{i,j} + u_{j,i})$ is the strain tensor, and τ_{ijk} is the higher-order stress. In this study, we adopt the gradient of the infinitesimal strain tensor $\varepsilon_{ijk} = \varepsilon_{ij,k}$ as the primal higher-order kinematic variable, which is referred as Mindlin's Form II. This choice simplifies the constitutive structure while retaining the ability to model the microstructural effect. There are classical viscosity and higher-order viscosity. Physically, the classical viscosity governs the dissipation due to the rate of local deformation, the higher-order viscosity captures the energy losses arising from the spatially heterogeneous evolution of the

microstructure. In complex materials such as biological tissues or polymers, this can represent the coarse-grained structural rearrangements, the friction associated with the relative motion between neighboring micro-constituents in regions of high strain gradients.

Then, the entropy inequality, i.e., Clausius–Duhem inequality, can be written as

$$\rho T\dot{s} - \rho r + Q_{i,i} - \frac{Q_i T_{,i}}{T} \geq 0, \tag{3}$$

where s is entropy per unit mass, T is temperature. Introducing the Helmholtz free energy per unit mass,

$$H = e - Ts, \tag{4}$$

and combining with Eqs. (1–3), lead to the inequality

$$\sigma_{ij}\dot{\varepsilon}_{ij} + \tau_{ijk}\dot{\varepsilon}_{ijk} - \rho\dot{H} - \rho s\dot{T} - \frac{Q_i T_{,i}}{T} \geq 0. \tag{5}$$

Working in small-strain linearized kinematics, the free energy per unit mass at current time depends on the current strain $\varepsilon(t)$, its spatial gradient $\varepsilon_{ijk}(t)$, and the past histories of these fields as well as the temperature history, i.e.,

$$\rho H(t) = \rho H_0 + F\Big\{[\varepsilon(\gamma),\ \nabla\varepsilon(\gamma),\ T(\gamma)]_{\gamma\leq t}\Big\}, \tag{6}$$

and it can be expanded into the polynomial representation:

$$\begin{aligned}\rho H(t) = \rho H_0 &+ \int_{-\infty}^{t} B_{ij}(t-\gamma)\ \dot{\varepsilon}_{ij}(\gamma)\mathrm{d}\gamma + \int_{-\infty}^{t} C_{ijk}(t-\gamma)\ \dot{\varepsilon}_{ijk}(\gamma)\mathrm{d}\gamma - \int_{-\infty}^{t} \beta(t-\gamma)\ \dot{T}(\gamma)\mathrm{d}\gamma \\ &+ \frac{1}{2}\int_{-\infty}^{t}\int_{-\infty}^{t} D_{ijkl}(t-\gamma,\ t-\vartheta)\ \dot{\varepsilon}_{ij}(\gamma)\ \dot{\varepsilon}_{kl}(\vartheta)\mathrm{d}\gamma\mathrm{d}\vartheta \\ &+ \frac{1}{2}\int_{-\infty}^{t}\int_{-\infty}^{t} H_{ijklmn}(t-\gamma,\ t-\vartheta)\ \dot{\varepsilon}_{ijk}(\gamma)\ \dot{\varepsilon}_{lmn}(\vartheta)\mathrm{d}\gamma\mathrm{d}\vartheta \\ &- \frac{1}{2}\int_{-\infty}^{t}\int_{-\infty}^{t} \psi(t-\gamma,\ t-\vartheta)\ \dot{T}(\gamma)\ \dot{T}(\vartheta)\mathrm{d}\gamma\mathrm{d}\vartheta + O(\varepsilon^3),\end{aligned} \tag{7}$$

where H_0 is the mean free energy, the integrating functions or mechanical properties $B_{ij}(t)$, $C_{ijk}(t)$, $\beta(t)$, $D_{ijkl}(t)$, $H_{ijklmn}(t)$, $\psi(t)$ are assumed to be continuous for $t \geq 0$ and vanish for $t < 0$. The terms of $O(\varepsilon^3)$ are neglected. The coupling terms between strain and strain gradient are neglected assuming the material is centrosymmetric, for which odd-order constitutive tensors vanish (Mindlin, 1964). Also, we assume the thermal coupling to the gradient field is negligible compared to the mechanical gradient effects for the isothermal or near-isothermal processes considered here.

In the following, the Stieltjes convolution symbol $*$, i.e.,

$$\phi * \mathrm{d}\varphi = \phi(t)\varphi(0) + \int_{-\infty}^{t} \phi(t-\zeta)\dot{\varphi}(\zeta)\mathrm{d}\zeta, \tag{8}$$

will be used. Substituting Eq. (7) into Eq. (5), one obtains

$$\begin{aligned}&\Big(-B_{ij}^0 - D_{ijkl}{}^*\mathrm{d}\varepsilon_{kl} + \sigma_{ij}\Big)\dot{\varepsilon}_{ij}(t) + \Big(-C_{ijk}^0 - H_{ijklmn}{}^*\mathrm{d}\varepsilon_{lmn} + \tau_{ijk}\Big)\dot{\varepsilon}_{ijk}(t) + \big(\beta^0 + \psi^*\mathrm{d}T - \rho s\big)\dot{T}(t) \\ &- \int_{-\infty}^{t} \frac{\partial B_{ij}}{\partial t}(t-\gamma)\ \dot{\varepsilon}_{ij}(\gamma)\mathrm{d}\gamma - \int_{-\infty}^{t} \frac{\partial C_{ijk}}{\partial t}(t-\gamma)\ \dot{\varepsilon}_{ijk}(\gamma)\mathrm{d}\gamma + \int_{-\infty}^{t} \frac{\partial \beta}{\partial t}(t-\gamma)\ \dot{T}(\gamma)\mathrm{d}\gamma \\ &- \frac{1}{2}\int_{-\infty}^{t}\int_{-\infty}^{t} \frac{\partial D_{ijkl}}{\partial t}(t-\gamma,\ t-\vartheta)\ \dot{\varepsilon}_{ij}(\gamma)\ \dot{\varepsilon}_{kl}(\vartheta)\mathrm{d}\gamma\mathrm{d}\vartheta \\ &- \frac{1}{2}\int_{-\infty}^{t}\int_{-\infty}^{t} \frac{\partial H_{ijklmn}}{\partial t}(t-\gamma,\ t-\vartheta)\ \dot{\varepsilon}_{ijk}(\gamma)\ \dot{\varepsilon}_{lmn}(\vartheta)\mathrm{d}\gamma\mathrm{d}\vartheta \\ &+ \frac{1}{2}\int_{-\infty}^{t}\int_{-\infty}^{t} \frac{\partial \psi}{\partial t}(t-\gamma,\ t-\vartheta)\ \dot{T}(\gamma)\ \dot{T}(\vartheta)\mathrm{d}\gamma\mathrm{d}\vartheta - \frac{Q_i T_{,i}}{T} \geq 0\ ,\end{aligned} \tag{9}$$

where $B_{ij}^0 = B_{ij}(0)$ is the initial stress, $C_{ijk}^0 = C_{ijk}(0)$ is the initial higher-order stress, and $\beta^0 = \beta(0)$ is the initial entropy. To satisfy the inequality for arbitrary rates, the coefficients multiplying these must give the constitutive identifications for the stress, the higher-order stress

$$\sigma_{ij} = B_{ij}^0 + D_{ijkl}{}^*\mathrm{d}\varepsilon_{kl}, \tag{10}$$

$$\tau_{ijk} = C_{ijk}^0 + H_{ijklmn}{}^*\mathrm{d}\varepsilon_{lmn}, \tag{11}$$

and the entropy

$$\rho s = \beta^0 + \psi^*\mathrm{d}T. \tag{12}$$

The fourth-order tensor D_{ijkl} and sixth-order tensor H_{ijklmn} can be related through the scale parameter, which we will derive in the below. The general framework derived above accounts for full thermomechanical coupling, which is ready for future multiphysics applications where thermal coupling is significant. The specific focus of this study is on the interaction between strain gradients and viscoelastic relaxation. In the remainder of this work, we assume isothermal conditions, i.e., $\dot{T} = 0$, $\nabla T = 0$. This simplifies the constitutive equations by removing thermal expansion and entropy flux terms, allowing us to isolate the mechanical size effects.

If we consider the initial state has no stress and higher-order stress, and write the deformation to volumetric strain and shape change, the viscoelastic stress–strain relation in Eq. (10) can be written as

$$\sigma_{ij} = \delta_{ij}\lambda * \mathrm{d}\varepsilon_{kk} + 2G * \mathrm{d}\varepsilon_{ij}, \tag{13}$$

where $G(t)$ and $\lambda(t)$ are relaxation functions. Similarly, the higher-order stress–strain gradient relation becomes

$$\tau_{ijk} = \delta_{ij}A_1 * \mathrm{d}\varepsilon_{ppk} + 2A_2 * \mathrm{d}\varepsilon_{ijk}, \tag{14}$$

where $A_1(t)$ and $A_2(t)$ are higher-order relaxation functions for isotropic solids. The relation between higher-order material parameters and the classical material parameters are defined as

$$A_1 = \lambda * \mathrm{d}c, \quad A_2 = G * \mathrm{d}c, \tag{15}$$

where $c = c(t)$ is the gradient viscoelastic parameter. This simplification allows our model to capture complex behaviors where the size effect evolves during the deformation process. This intrinsic material parameter c, governs the magnitude of the strain gradient effect and endows the material model with size-dependency. A general isotropic strain gradient theory, requires 18 independent constitutive parameters. Calibrating such a complex model is often experimentally intractable. Therefore, we adopt a simplified strain gradient formulation (Askes and Aifantis, 2011; Gao and Park, 2007). It reduces the higher-order constitutive tensor to a dependence on a single internal scale parameter. While spatially simplified, we retain temporal generality by allowing this gradient parameter to evolve over time, thereby capturing the distinct relaxation kinetics of the microstructure relative to the bulk continuum.

Then, the governing equilibrium equations should be derived. At current time, the virtual work statement can be constructed as

$$\int_V (\sigma_{ij}\delta\varepsilon_{ij} + \tau_{ijk}\delta\varepsilon_{ijk})\mathrm{d}V = \int_V (f_i\delta u_i)\mathrm{d}V + \int_S (p_i\delta u_i + q_i D\delta u_i)\mathrm{d}S, \tag{16}$$

where $f_i(t)$,$p_i(t)$ and $q_i(t)$ are the body force, the traction and higher-order traction, respectively. $Du_i = n_l\partial_l u_i$ is the normal derivative of the displacement on the boundary with n_l being the components of the outward unit normal vector on $\mathrm{d}S$. Applying the divergence theorem and integration by parts, one obtains

$$\begin{aligned} -\int_V (\sigma_{ij,j} - \tau_{ijk,jk})\delta u_i \mathrm{d}V + \int_S [(\sigma_{ij} - \tau_{ijk,k})n_j\delta u_i + (\tau_{ijk}n_k)\delta u_{i,j}]\mathrm{d}S \\ = \int_V (f_i\delta u_i)\mathrm{d}V + \int_S (p_i\delta u_i + q_i D\delta u_i)\mathrm{d}S. \end{aligned} \tag{17}$$

The equilibrium equation can be obtained

$$\left(\sigma_{ij} - \tau_{ijk,k}\right)_{,j} + f_i = 0, \tag{18}$$

together with the corresponding natural boundary conditions

$$\begin{aligned} (\sigma_{ij} - \tau_{ijk,k})n_j - D_j(n_k\tau_{ijk}) + (D_l n_l)\tau_{ijk}n_k n_j = \widehat{p}_i, \\ \tau_{ijk}n_j n_k = \widehat{q}_i, \end{aligned} \tag{19}$$

where $D_j(\bullet) = (\delta_{jl} - n_j n_l)\partial_l(\bullet)$ is the surface gradient operator.

3. Numerical implementation of strain gradient viscoelasticity

The direct finite element implementation of the SGV equilibrium equation Eq. (18) is challenging. The strain gradient term involves the second-order spatial derivatives of the displacement field, which would necessitate the use of C^1-continuous finite elements. To circumvent this continuity requirement, we adopt a mixed-field finite element formulation, which introduces additional node degrees of freedoms to lower the order of the derivatives in the variational form. Moreover, we also build a PINN framework for the strain gradient viscoelasticity, which can decrease the requirement of mixed formulation.

3.1. Governing equations and mixed variational principle

For the strain gradient elasticity, the mixed-type finite element formulation has been developed (Amanatidou and Aravas, 2002), including the three-dimensional one (Zybell et al., 2012). In the following, we derive the mixed-type finite element formulation of SGV theory. The additional degree of freedoms ϕ_{ij} and ρ_{ij} are included in this mixed-type formulation. They satisfy

$$\phi_{ij} \approx \varepsilon_{ij}, \quad \rho_{ij} \approx \tau_{ijk,k}. \tag{20}$$

Therefore, the strain gradient is relaxed as η_{ijk}, i.e.,

$$\eta_{ijk} = \phi_{ij,k}, \tag{21}$$

and the higher-order stress in the finite element calculation becomes

$$\tau_{ijk} = \delta_{ij}\delta_{jk}A_1 * \mathrm{d}\eta_{ppp} + 2A_2 * \mathrm{d}\eta_{ijk}. \tag{22}$$

Then, the equilibrium equation Eq. (18) is written as

$$\left(\sigma_{ij} - \rho_{ij}\right)_{,j} + f_i \approx 0. \tag{23}$$

Therefore, the problem can be formed as a coupled system of equations: the equilibrium equation Eq. (23) and the kinematic constraints Eq. (20). The resulting variational statement at current time requires that:

$$\begin{aligned} &\int_V (\phi_{ij} - \varepsilon_{ij})\delta\rho_{ij}\mathrm{d}V = 0, \\ &\int_V (\tau_{ijk,k} - \rho_{ij})\delta\phi_{ij}\mathrm{d}V = 0, \\ &\int_V \left[\left(\sigma_{ij} - \rho_{ij}\right)_{,j} + f_i\right]\delta u_i \mathrm{d}V = 0. \end{aligned} \tag{24}$$

Applying the Gauss divergence theorem and integrating by parts, Eq. (24) can be rewritten as

$$\begin{aligned} &\int_V (\phi_{ij} - \varepsilon_{ij})\delta\rho_{ij}\mathrm{d}V = 0, \\ &\int_V (\tau_{ijk}\delta\eta_{ijk} + \rho_{ij}\delta\phi_{ij})\mathrm{d}V = \int_S s_{ij}\delta\phi_{ij}\mathrm{d}S, \\ &\int_V (\sigma_{ij} - \rho_{ij})\delta\varepsilon_{ij}\mathrm{d}V = \int_V f_i\delta u_i\mathrm{d}V + \int_S p_i\delta u_i\mathrm{d}S, \end{aligned} \tag{25}$$

where p_i denotes the surface traction, s_{ij} corresponds to the higher-order traction. Eq. (25) represents the general variational principle including external work on the boundary. In this study, we assume the higher-order boundaries are free of external double tractions. Consequently, the boundary integral $\int_S s_{ij}\delta\phi_{ij}\mathrm{d}S$ vanishes. Dropping this term and neglecting body forces, we arrive at the simplified weak form used for the finite element implementation:

$$\delta\left[\int_V \left(\frac{1}{2}\sigma_{ij}\varepsilon_{ij} + \frac{1}{2}\tau_{ijk}\eta_{ijk} + (\phi_{ij} - \varepsilon_{ij})\rho_{ij}\right)\mathrm{d}V\right] = \int_S p_i\delta u_i\mathrm{d}S. \tag{26}$$

It can be found that the mixed-type finite element formulation only contains first-order derivatives of the independent variables (u_i, ϕ_{ij}, ρ_{ij}) and balances them in the weak form of the variational formulation.

3.2. Variational formulation and time integration

The Stieltjes convolution Eqs. (13) and (14), representing the memory of materials, is computationally expensive to evaluate directly as it requires storing the entire strain history. This challenge can be overcome using the recurrence formula, which is possible when the relaxation functions are represented by a Prony series. For demonstration, we apply the three-parameter model (a spring placed in parallel with a Maxwell element, as shown in Fig. 1a for 1D case) to describe the viscoelastic behavior. This model is represented by a fourth-order

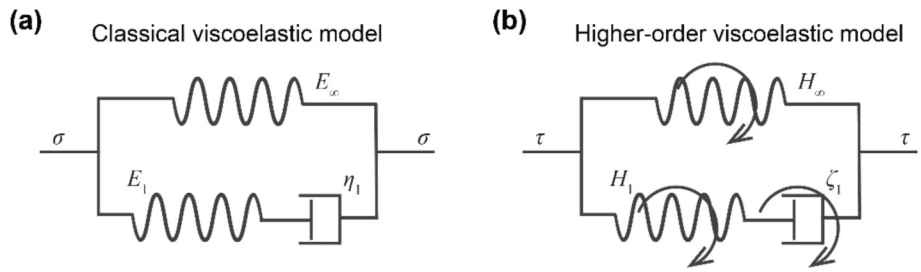

Fig. 1. The three-parameter viscoelastic models. (a) classical and (b) higher-order viscoelastic models. The relaxation time of the classical viscoelastic model is $\tau_g = \eta_1/E_1$, and the relaxation time is $\kappa_g = \zeta_1/H_1$ for the higher-order viscoelastic model.

relaxation tensor $D_{ijkl}(t)$.

Considering the isotropic material, the Cauchy stress expressed in Eq. (13) can be separated into two parts

$$\sigma_{ij}(t) = D^{\infty}_{ijkl}\varepsilon_{kl}(t) + \sigma^{r}_{ij}, \tag{27}$$

where σ^{r}_{ij} is the stress acts on the Maxwell branch, i.e.,

$$\dot{\sigma}^{r}_{ij}(t) + \frac{\sigma^{r}_{ij}(t)}{\tau_g} = D^{1}_{ijkl}\dot{\varepsilon}_{kl}(t), \tag{28}$$

D^{∞}_{ijkl} is the long-term stiffness tensor, D^{1}_{ijkl} is the stiffness tensor of the Maxwell branch, and τ_g is the relaxation time. In the following isotropic beams, the relaxation time is calculated as the viscous parameter divided by the elastic modulus. Using the backward Euler method, Eq. (28) can be written as

$$\frac{\sigma^{r}_{ij,n} - \sigma^{r}_{ij,n-1}}{\Delta t} + \frac{\sigma^{r}_{ij,n}}{\tau_g} = D^{1}_{ijkl}\frac{\varepsilon_{kl,n} - \varepsilon_{kl,n-1}}{\Delta t} \tag{29}$$

where the time increment is Δt and current time $t_n = t_{n-1} + \Delta t$. We should select a time step Δt less than the relaxation time to ensure the physical accuracy of the relaxation. The subscript n and $n-1$ denote the current step and previous step, respectively. Then, the recurrence formula of stress σ^{r}_{ij} can be obtained

$$\sigma^{r}_{ij,n} = \frac{1}{1 + \Delta t/\tau_g}\left[\sigma^{r}_{ij,n-1} + D^{1}_{ijkl}\left(\varepsilon_{kl,n} - \varepsilon_{kl,n-1}\right)\right]. \tag{30}$$

Therefore, the stress at the current time step can be expressed as

$$\sigma_{ij,n} = \left(D^{\infty}_{ijkl} + \frac{D^{1}_{ijkl}}{1 + \Delta t/\tau_g}\right)\varepsilon_{kl,n} + \frac{\sigma^{r}_{ij,n-1} - D^{1}_{ijkl}\varepsilon_{kl,n-1}}{1 + \Delta t/\tau_g} = D^{\text{eff}}_{ijkl}\varepsilon_{kl,n} + \sigma^{\text{hist}}_{ij}, \tag{31}$$

where D^{eff}_{ijkl} is an effective stiffness tensor, and $\sigma^{\text{hist}}_{ij} = \left(\sigma^{r}_{ij,n-1} - D^{1}_{ijkl}\varepsilon_{kl,n-1}\right)/\left(1 + \Delta t/\tau_g\right)$ denotes the history effect contribution of the stress. Similarly, the recurrence formula of higher-order stress is derived as

$$\begin{aligned}\tau_{ijk,n} &= \left(H^{\infty}_{ijkabc} + \frac{H^{1}_{ijkabc}}{1 + \Delta t/\kappa_g}\right)\varepsilon^{h}_{abc,n} + \frac{\tau^{r}_{ijk,n-1} - H^{1}_{ijkabc}\varepsilon^{h}_{abc,n-1}}{1 + \Delta t/\kappa_g} \\ &= H^{\text{eff}}_{ijkabc}\varepsilon^{h}_{abc,n} + \tau^{\text{hist}}_{ijk},\end{aligned} \tag{32}$$

where H^{∞}_{ijkabc} and H^{1}_{ijkabc} are higher-order stiffness tensors obtained from Fig. 1b, τ^{r}_{ijk} is the higher-order stress acts on the higher-order Maxwell unit, H^{eff}_{ijkabc} is an effective higher-order elastic modulus, $\tau^{\text{hist}}_{ijk} = \left(\tau^{r}_{ijk,n-1} - H^{1}_{ijkabc}\varepsilon^{h}_{abc,n-1}\right)/\left(1 + \Delta t/\kappa_g\right)$ denotes the history effect contribution of the higher-order stress, and κ_g is the microstructural relaxation time. In this study, it is assumed that the conventional shear and bulk moduli are governed by the same relaxation time τ_g, and the higher-order moduli share a single uniform relaxation time κ_g. For certain physical conditions, the time-dependent behavior in the shear and bulk directions may differ. Relaxation times can be modified according to experimental measurements.

Furthermore, the principle of virtual work at current time step is

$$\int_{\Omega}\left(\sigma_{ij,n}\delta\varepsilon_{ij} + \tau_{ijk,n}\delta\varepsilon^{h}_{ijk,n}\right)\mathrm{d}V = \int_{\Omega} f_i\delta u_i\mathrm{d}V + \int_{\partial\Omega_T}\left[p_i\delta u_i + q_i D(\delta u_i)\right]\mathrm{d}S. \tag{33}$$

Substituting Eqs. (31) and (32) in Eq. (33), and ignoring the body force, the weak form becomes

$$\begin{aligned}&\int_{\Omega}\left[\left(D^{\text{eff}}_{ijkl}\varepsilon_{kl,n} + \sigma^{\text{hist}}_{ij}\right)\delta\varepsilon_{ij} + \left(H^{\text{eff}}_{ijkabc}\varepsilon^{h}_{abc,n} + \tau^{\text{hist}}_{ijk}\right)\delta\varepsilon^{h}_{ijk}\right]\mathrm{d}V \\ &\quad = \int_{\partial\Omega_T}\left[p_i\delta u_i + q_i D(\delta u_i)\right]\mathrm{d}S.\end{aligned} \tag{34}$$

Then, a functional for the finite element implementation at each time step can be set as

$$\begin{aligned}\Pi_0(t) = \frac{1}{2}\int_{\Omega}\left(\varepsilon_{ij}D^{\text{eff}}_{ijkl}\varepsilon_{kl} + \varepsilon^{h}_{ijk}H^{\text{eff}}_{ijkabc}\varepsilon^{h}_{abc}\right)\mathrm{d}V + \int_{\Omega}\left(\sigma^{\text{hist}}_{ij}\varepsilon_{ij}\right. \\ \left. + \tau^{\text{hist}}_{ijk}\varepsilon^{h}_{ijk}\right)\mathrm{d}V - \int_{\partial\Omega_T}(p_i u_i + q_i D u_i)\mathrm{d}S.\end{aligned} \tag{35}$$

It should be noted that the history effects (i.e., $\sigma^{\text{hist}}_{ij}$ and τ^{hist}_{ijk}) are known because they are calculated from the previous step state.

For the mixed-type element, additional degree of freedoms ϕ and ρ are included. According to Eqs. (25) and (26), the functional should also include

$$\Pi_{\text{mix}}(t) = \int_{\Omega}\rho_{ij}\,\phi_{ij}\,\mathrm{d}V - \int_{\Omega}\rho_{ij}\,\varepsilon_{ij}\,\mathrm{d}V. \tag{36}$$

Therefore, the final formula of the functional for the mixed-type element is

$$\Pi = \Pi_0 + \Pi_{\text{mix}}. \tag{37}$$

Then, the finite element model based on the above formulations can be accomplished using the open source FEniCS project (Alnæs et al., 2014; Scroggs et al., 2022a, 2022b; Baratta et al., 2025), where the domain is discretized into a mesh of finite elements and the problem is solved by finding the minimum value of the functional.

3.3. Physics-informed neural network implementation

This section details the physics-informed neural networks (PINNs) framework developed to solve the initial boundary value problem of strain gradient viscoelasticity. We adapt a variational energy-based PINNs, also known as the deep energy method (Bai et al., 2025; Samaniego et al., 2020; Zhuang et al., 2021), by introducing the time integration and the higher-order boundary conditions. We integrate this framework with the recurrence, incremental time-stepping procedure derived in Section 3.2. The overall workflow, illustrated schematically

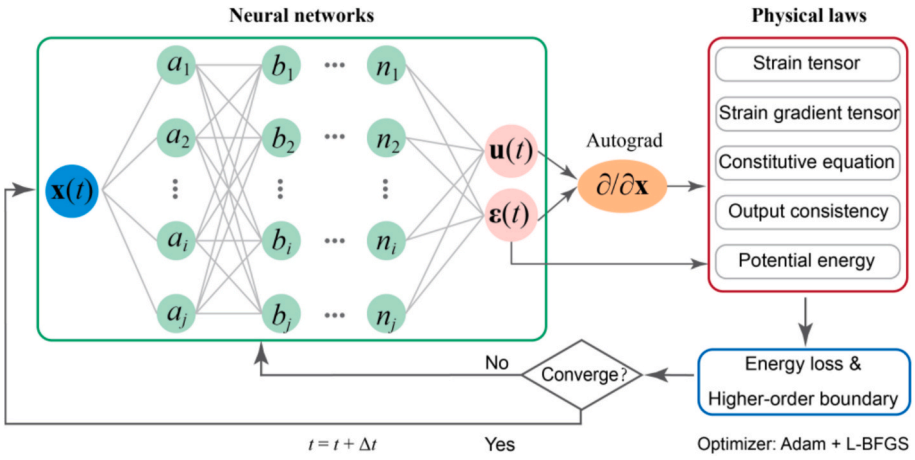

Fig. 2. Schematic of the physics-informed neural network for strain gradient viscoelasticity. At each time step, the network takes coordinates $\boldsymbol{x}$ as input and outputs displacement and an additional strain field. These are used to compute the potential energy and other loss terms, which are minimized to update the trainable parameters of networks.

in Fig. 2, solves a pseudo-elastic problem at each time step to capture the full time-dependent material response.

3.3.1. Formulation of PINNs architecture

In the energy-based PINNs for viscoelasticity, the core idea is to approximate the displacement field $\boldsymbol{u}_{NN}$ at discrete time points t_i ($i = 0, 1, \ldots, n$) using the neural networks (Abueidda et al., 2022; Lin et al., 2026). The network takes spatial coordinates $\boldsymbol{x}$ as input and outputs the displacement vector at that specific time step:

$$\boldsymbol{u}_{NN}(\boldsymbol{x}, t_i) = N_u(\boldsymbol{x}; \boldsymbol{w}_u), \tag{38}$$

where $N_u(\cdot)$ is the neural network mapping with trainable parameters $\boldsymbol{w}_u$ at time step t_i, and $\boldsymbol{x}$ is the position of material points. A key advantage of this framework is that spatial derivatives of $\boldsymbol{u}_{NN}$ required for computing strain and strain gradient can be calculated analytically and efficiently using automatic differentiation.

While automatic differentiation can compute the second-order derivatives required for SGV, directly optimizing a loss function involving such terms can be challenging, mainly because the loss term with higher-order derivatives tends to be optimized to zero. Therefore, we also use the "mixed-type" idea in the PINNs framework, but not as strict as in the finite element. The neural networks also output an additional strain field, as shown in Fig. 2, which is used to calculated the strain gradient, i.e.,

$$\begin{aligned} \left\{\boldsymbol{u}_{NN}(\boldsymbol{x}, t_i),\ \varepsilon^d_{ij}(\boldsymbol{x}, t_i)\right\} &= N(\boldsymbol{x}; \boldsymbol{w}), \\ \varepsilon_{ijk} &= \varepsilon^d_{ij,k}. \end{aligned} \tag{39}$$

The directed output strains are enforced by a regularization term, $\left|\varepsilon_{ij} - \varepsilon^d_{ij}\right|$, where ε_{ij} is calculated based on the output displacement $\boldsymbol{u}_{NN}$. In practice, the neural network architecture, i.e. the number of hidden layers and the number of nodes per layer, is flexible and can be adjusted based on the PDE complexity (Li et al., 2024; Nguyen-Thanh et al., 2020). In this study, the neural network includes three hidden layers and each layer contains forty neurons, which is used to output the displacement and directed strain ε^d_{ij}. The tanh activation function is used in this study.

The network parameters $\boldsymbol{w}$ are optimized by minimizing a composite loss function that encodes the physics of the system. This loss function is constructed from three primary components. The first component is the potential energy. Under the same condition of vanishing external higher-order tractions, the potential energy at each time step can be written as

$$\begin{aligned} \Pi_{\mathrm{PINN}} = \frac{1}{2}\int_\Omega \left(\varepsilon_{ij} D^{\mathrm{eff}}_{ijkl} \varepsilon_{kl} + \varepsilon^h_{ijk} H^{\mathrm{eff}}_{ijkabc} \varepsilon^h_{abc}\right) \mathrm{d}V + \int_\Omega \left(\sigma^{\mathrm{hist}}_{ij} \varepsilon_{ij} \right. \\ \left. + \tau^{\mathrm{hist}}_{ijk} \varepsilon^h_{ijk}\right) \mathrm{d}V - \int_S p_i u_i \ \mathrm{d}S, \end{aligned} \tag{40}$$

according to the recurrence time-stepping formula derived in the previous section, i.e., Eqs. (27–35). This recurrence update allows for the efficient computation of the history-dependent stress without storing the full strain history. The second component is the kinematic consistency. This penalty term enforces that the additional strain output $\boldsymbol{\varepsilon}^d$ is consistent with the true strain computed from the displacement output $\boldsymbol{u}_{NN}$. It is formulated as a mean squared error:

$$L_{\mathrm{con}} = \beta \left\| \frac{1}{2}\left[\Delta \boldsymbol{u}_{NN} + (\Delta \boldsymbol{u}_{NN})^T\right] - \boldsymbol{\varepsilon}^d \right\|^2, \tag{41}$$

where the penalty factor β is chosen as 5e4 in this study. Furthermore, in the strain gradient theory, there are higher-order boundary conditions. We use the neural network and the automatic differentiation function to obtain the value on the specific boundary, and make them are satisfied with the boundary conditions, i.e.,

$$L_{\mathrm{hbc}} = \alpha_i (b_{\mathrm{bc\text{-}i}} - \widehat{b}_i)^2 \tag{42}$$

where $b_{\mathrm{bc\text{-}i}}$ is the value on the boundary i that obtained from the neural network, $\widehat{b}_i$ is the boundary constraint value, and α_i is the penalty factor.

Therefore, the network parameters $\boldsymbol{w}$ can be found by minimizing the loss function:

$$\begin{aligned} \boldsymbol{w} &= \mathrm{argmin}_{\boldsymbol{u}, \varepsilon_d} \mathscr{L} \\ &= \mathrm{argmin}_{\boldsymbol{u}, \varepsilon_d} \left[\Pi_{\mathrm{PINN}}(\boldsymbol{u},\ \partial \boldsymbol{u}/\partial \boldsymbol{x}\ , \partial \boldsymbol{\varepsilon}^d/\partial \boldsymbol{x},\ G,\ \lambda,\ A_1,\ A_2) + L_{\mathrm{con}} + L_{\mathrm{hbc}}\right], \end{aligned} \tag{43}$$

using the gradient descent method (Adam (Kingma and Ba, 2017) first and then L-BFGS in this study). Once converged, the displacement field $\boldsymbol{u}(\boldsymbol{x}, t_i)$ and the directed strain are $\boldsymbol{\varepsilon}^d$ obtained, and then the corresponding strain tensor, strain gradient tensor can be obtained. The stress tensor, higher-order stress tensor, and the history-dependent items can be calculated and updated. Then, the simulation advances to the next time step t_{i+1} with increment Δt, and the time increment should be smaller than the relaxation time to avoid inaccurate results. This incremental process allows simulation of the full time-dependent viscoelastic response.

3.3.2. Implementation of boundary conditions

The traction boundary conditions are naturally incorporated through the external work term in the potential energy functional, and the higher-order boundary conditions are included explicitly in the loss function as describe by Eq. (42). For classical displacement boundary conditions, there are two main approaches to impose within energy-based PINNs: the soft and hard ways. The soft way means adding a penalty term to the loss function that penalizes deviations from the prescribed displacement (Bai et al., 2025; Barrett and Elliott, 1986), like we did for the higher-order boundary conditions. We use the hard enforcement to impose the displacement boundary conditions. This method imposes the boundary conditions exactly without adding penalty terms to the loss function, which can be beneficial for training convergence (Bai et al., 2025). $\boldsymbol{u}(\boldsymbol{x},t)=\widehat{\boldsymbol{u}}(\boldsymbol{x},t)$ for $\boldsymbol{x}\in\Gamma(\boldsymbol{x})$ can satisfy by modifying the displacement filed output by PINNs through distance functions:

$$\boldsymbol{u}(\boldsymbol{x},t_i)=N(\boldsymbol{x};\boldsymbol{w})\odot b(\boldsymbol{x})+\widehat{\boldsymbol{u}}(\boldsymbol{x},t_i), \tag{44}$$

where $\odot$ means the element-wise product, and b is an approximate distance function to the boundary $\Gamma(\boldsymbol{x})$ which denotes the shortest distance of a point $\boldsymbol{x}$ to the essential boundary. The distance function is non-negative and $b(\boldsymbol{x})=0$ only when $\boldsymbol{x}\in\Gamma(\boldsymbol{x})$. For the simple geometry with explicit boundaries, the method in Eq. (44) is easy to implement. For complex geometries, the boundary conditions can imposed using distance functions (Sukumar and Srivastava, 2022). Different methods to impose displacement boundary conditions on PINNs have been listed and compared by Berrone et al. (Berrone et al., 2023).

4. Analytical and numerical results

This section presents a comprehensive analysis of cantilever and simply-supported beam bending problems to validate the proposed numerical frameworks and investigate the physical predictions of the SGV theory. The slenderness of the two-dimensional (2D) beam used for numerical simulation is set as $L=10\,h$, where L is the beam length, and h is the beam height. The material parameters are $E_\infty=E_1$ and the Poisson's ratio $\nu=0$ in the numerical analysis in order to more correspond to the one-dimensional analytical Bernoulli-Euler beam solution. The beam is subjected to the external pressure $q(t)=q_0\vartheta(t)$ on the top surface, where $\vartheta(t)$ denotes the Heaviside function. The material parameters used for the benchmark problems in this section are chosen for illustrative purposes to verify the numerical implementation. They represent a generic viscoelastic solid with properties selected to highlight the interaction between the internal length scale and the structural dimensions.

For the finite element implementation of the beam, the FEniCSx-0.7 is used with first-order quadrilateral elements. The element size was set to $h/10$ unless otherwise specified, resulting in a 100×10 mesh. Each node of this 2D model has eight degrees of freedom corresponding to the mixed-field variables, i.e., $\left(u_x,\,u_y,\,\phi_{xx},\,\phi_{yy},\,\phi_{xy},\,\rho_{xx},\,\rho_{yy},\,\rho_{xy}\right)$. The boundary conditions, including the higher-order ones (Eqs. (A7) and (B4)), can be applied. Then, the functional Eq. (37) is solved using the Newton–Raphson method. For the PINNs implementation, the density of collocation points was selected to match the nodal resolution of the finite element mesh to ensure a consistent comparison. The higher-order boundary conditions of beams are enforced softly via a penalty term. As described by Eqs. (A7) and (B4) corresponding to cantilever beam and simply-supported beam respectively, the higher-order differential is obtained using the automatic differential function of Pytorch. The training of PINNs constitutes a multi-objective optimization task where the total loss is a weighted sum of the governing equation residuals and boundary condition errors, as descript by Eq. (43). The selection of loss weights follows a systematic approach based on balancing the gradient magnitudes. In this study, the higher-order boundary conditions can be applied according to Eq. (42), with the penalty $\alpha_i=500$. The penalty factor β in Eq. (41) is chosen as $\beta=5\mathrm{e}4$.

4.1. Analytical formulation for a Bernoulli-Euler beam

The bending of a beam is selected as a representative model in this study because it is a canonical problem in solid mechanics. Its well-understood kinematics provide a robust foundation upon which the more complex interactions of strain gradient and viscoelastic effects can be clearly investigated and isolated. A Bernoulli-Euler beam of length L, width b, and height h, is considered. The x axis is colinear with the centerline of the beam, and y, z axes are within the cross-section. x, y, z are enumerated as 1, 2, 3. The displacement components are given as

$$u_1=-z\phi(x,t),u_2=0,u_3=w(x,t), \tag{45}$$

where $\phi(x,t)$ denotes the rotation angle of the beam cross-sections around $y-$axis. For a slender beam, the transverse shear is always ignored, that is $\phi\simeq \mathrm{d}w/\mathrm{d}x=w'$.

Then, the non-vanishing component of strain is

$$\varepsilon_{11}=-zw'', \tag{46}$$

and the strain gradient components are

$$\varepsilon_{111}=-zw''',\varepsilon_{113}=-w''. \tag{47}$$

The stress σ_{11} can be obtained by neglecting the Poisson effect, i.e.,

$$\sigma_{11}=E*\mathrm{d}\varepsilon_{11}=-zE*\mathrm{d}w'', \tag{48}$$

and higher-order stress τ_{111}, τ_{113} are

$$\begin{aligned}\tau_{111}&=A*\mathrm{d}\varepsilon_{111}=-zA*\mathrm{d}w'''=-zE*\mathrm{d}c*\mathrm{d}w''',\\ \tau_{113}&=A*\mathrm{d}\varepsilon_{113}=-A*\mathrm{d}w''=-E*\mathrm{d}c*\mathrm{d}w''.\end{aligned} \tag{49}$$

We define the total stress as $\mu_{ij}=\sigma_{ij}-\tau_{ijk,k}$. The total normal stress in x direction can be obtained

$$\mu_{11}=\sigma_{11}-\tau_{111,1}-\tau_{113,3}=-zE*\mathrm{d}w''+zE*\mathrm{d}c*\mathrm{d}w^{\mathrm{IV}}. \tag{50}$$

Conditions of equilibrium require that the resultant of the internal forces on the cross-section should be zero, and the moment equals the bending moment, that is,

$$\int_S\mu_{11}\mathrm{d}S=0, \tag{51}$$

$$\int_S(z\mu_{11}+\tau_{113})\mathrm{d}S=-M, \tag{52}$$

with

$$\frac{\mathrm{d}M}{\mathrm{d}x}=Q,\quad\frac{\mathrm{d}Q}{\mathrm{d}x}=q(x). \tag{53}$$

Substituting Eqs. (49) and (50) into Eqs. (51) and (52), one obtains

$$\int_S\left(zE*\mathrm{d}w''+zA*\mathrm{d}w^{\mathrm{IV}}\right)\mathrm{d}S=\left(E*\mathrm{d}w''+A*\mathrm{d}w^{\mathrm{IV}}\right)\int_S z\mathrm{d}S=0, \tag{54}$$

$$\begin{aligned}&\int_S\left(-z^2E*\mathrm{d}w''+z^2A*\mathrm{d}w^{\mathrm{IV}}-A*\mathrm{d}w''\right)\mathrm{d}S\\ &=\left(-E*\mathrm{d}w''+A*\mathrm{d}w^{\mathrm{IV}}\right)\int_S z^2\mathrm{d}S-\int_S(A*\mathrm{d}w'')\mathrm{d}S=-M.\end{aligned} \tag{55}$$

Because the coordinate origin is the cross-section center, i.e., $\int_S z\mathrm{d}S=0$, Eq. (54) is naturally satisfied. With $\int_S z^2\mathrm{d}S=I$ denoting the moment of inertia, Eq. (55) becomes

$$IE*\mathrm{d}w''+aE*\mathrm{d}c*\mathrm{d}w''-IE*\mathrm{d}c*\mathrm{d}w^{\mathrm{IV}}=M, \tag{56}$$

where a is the area of the cross-section. From Eqs. (53) and (56), the governing equation of Bernoulli-Euler beam in bending can be described as

$$(IE + aE * \mathrm{d}c) * \mathrm{d}w^{\mathrm{IV}} - IE * \mathrm{d}c * \mathrm{d}w^{\mathrm{VI}} = q. \tag{57}$$

The governing equation and boundary conditions can also be derived using the variational principle, as in the paper of Lin and Wei (Lin and Wei, 2020).The boundary conditions at $x = 0,\ L$ are

$$\begin{aligned} M_l' - M_h'' &= Q_s \text{ or } w = w_0, \\ M_l - M_h' + M_s &= 0 \text{ or } w' = w_0', \\ M_h + M_{hs} &= 0 \text{ or } w'' = w_0'', \end{aligned} \tag{58}$$

where $Q_s(t)$ is the boundary shear force, $M_s(t)$ and $M_{hs}(t)$ are the boundary moment and higher-order moment. The moment M_l is

$$M_l = (IE + aA) * \mathrm{d}w'' = (IE + aE * \mathrm{d}c) * \mathrm{d}w'', \tag{59}$$

and the higher-order moment M_h is

$$M_h = M_h(t) = IA * \mathrm{d}w''' = IE * \mathrm{d}c * \mathrm{d}w'''. \tag{60}$$

Eq. (57) includes the effect of the deformation history, which is described by the Stieltjes integral. Therefore, the viscosity exhibited during beam bending can be explained. If the viscosity is not considered, the governing equation and boundary conditions will reduce to the classical gradient elastic Bernoulli-Euler beam model (Lurie and Solyaev, 2018)

$$(EI + cEa)w^{\mathrm{IV}} - IcEw^{\mathrm{VI}} = q. \tag{61}$$

The governing equilibrium equation for Bernoulli-Euler beam bending has been given. The solutions for different types of beams bending problems can be obtained by combining them with the appropriate boundary conditions. The theoretical solution of the cantilever beam and simply-supported beam can be found in Appendices A and B.

In the SGV theory, the gradient parameter is time-dependent and has a finite value, as in Eq. (A12). At $t = 0$, the gradient parameter is $c|_{t=0} = c_e$. For $t \to \infty$, the gradient parameter is $c|_{t\to\infty} = c_e$. c_e is an elastic gradient parameter which links the classical moduli and higher-order moduli (Appendix A). The relation between the gradient parameter and the ratio κ_g/τ_g has been discussed in the previous work (Lin and Wei, 2020). Furthermore, the higher-order viscosity related to the micro-/

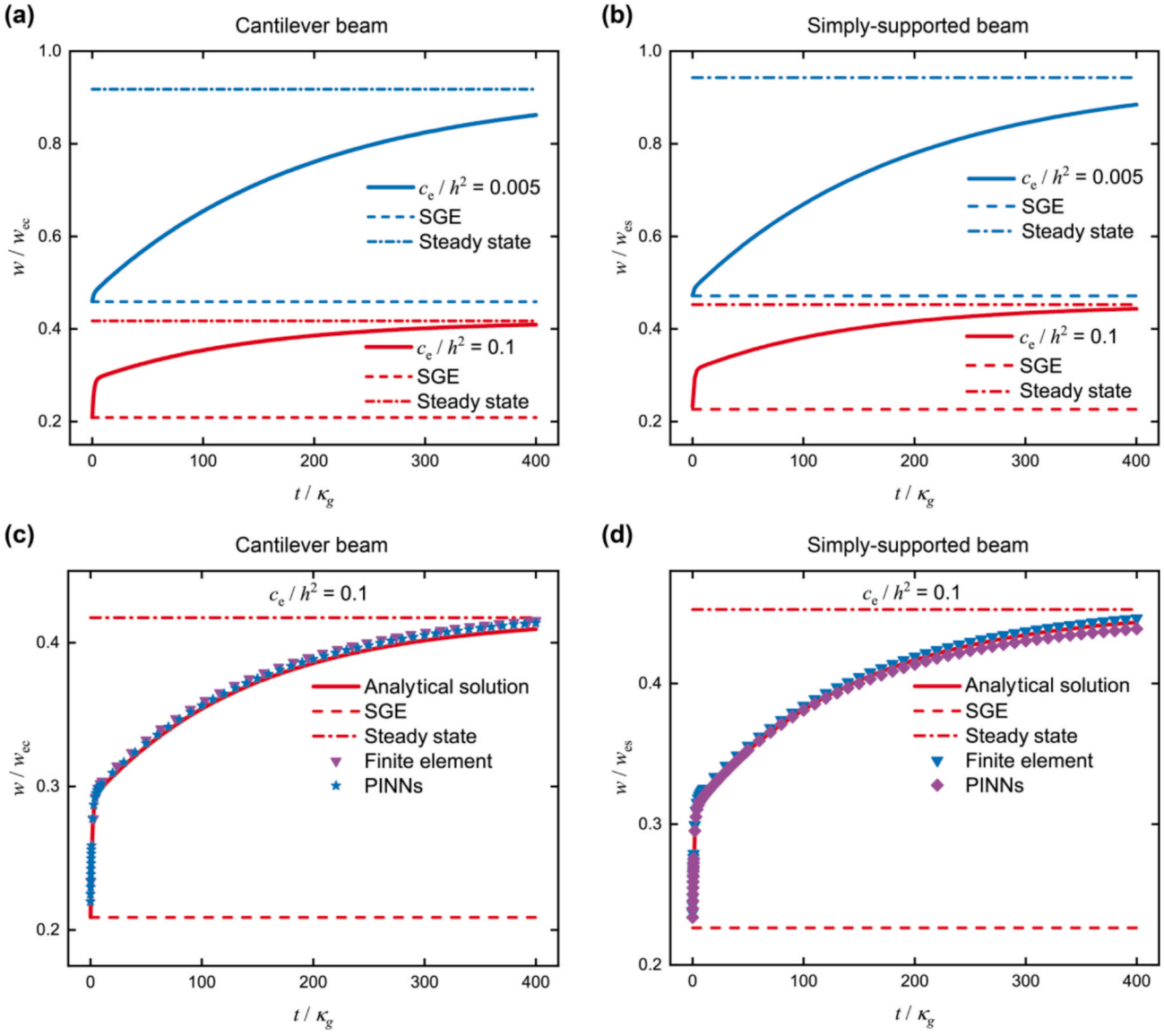

Fig. 3. Dimensionless deflection over time of beams, with $\kappa_g/\tau_g = 0.01$, $E_\infty = E_1$. *SGE* denotes the results of the classical strain gradient elasticity with the elastic modulus $E_0 = E_\infty + E_1$, and *Stead state* denotes the result at $t \to \infty$. (a) Deflection at the free end of the cantilever beam for different gradient parameters. (b) Deflection at the center of the simply-supported beam over time for different gradient parameters. (c-d) Deflections of the cantilever beam and simply-supported beam calculated using three methods: theoretical analysis, finite element method and PINNs, respectively.

nano-structure evolution should be more rate-sensitive than classical viscosity. Therefore, the classical relaxation time should be much larger than the higher-order relaxation time, that is $\tau_g > \kappa_g$. In this study, the relaxation times are set as $\kappa_g < \tau_g$.

4.2. Time-dependent deformation of the beam

The time-dependent beam deflection is concerned firstly. Fig. 3 shows the relation between dimensionless deflection and time for different dimensionless gradient parameters, where $w_{ec} = q_0 L^4/(8E_\infty I)$ and $w_{es} = 5qL^4/(384E_\infty I)$ are the maximum deflections of the cantilever beam and simply-supported beam in the classical elasticity, respectively. *SGE* in the legend denotes the classical strain gradient elasticity results with modulus $E_0 = E_\infty + E_1$, and gradient parameter $c_0 = c_e$. In Fig. 3a-b, the dimensionless deflection increases with time, i.e., the material undergoes creep or a gradual softening of the material. In the initial phase, the deflection ascends sharply with time because the microstructural viscoelasticity relaxes quickly. When the time continues to increase, the dimensionless deflection tends to a steady-state value. The results of gradient viscoelastic deflection are always larger than the SGE result, but smaller than the classical elastic result (the dimensionless deflection always smaller than 1.). There is a strengthening effect of our theory compared with the traditional elasticity theory, but a time-dependent softening evolution compared with the classical strain gradient elasticity.

Furthermore, we validate our finite element and PINNs implementations against the derived analytical solution. Fig. 3c-d plots the maximum beam deflection over time for both cantilever and simply-supported cases. An excellent agreement is observed among all three methods, confirming that both the mixed-field finite element and the PINNs accurately capture the SGV response and its temporal evolution. This indicates that these two numerical methods may provide solution for complex structural problems in the future.

For the recurrence method implemented in PINNs, there may be some errors, as shown in Fig. 4. There are two main reasons: first, the PINNs approximates the solution by minimizing a composite loss function. This loss function represents a multi-objective optimization problem, balancing the governing physical laws (e.g., potential energy), higher-order boundary conditions, and other constraints (constant

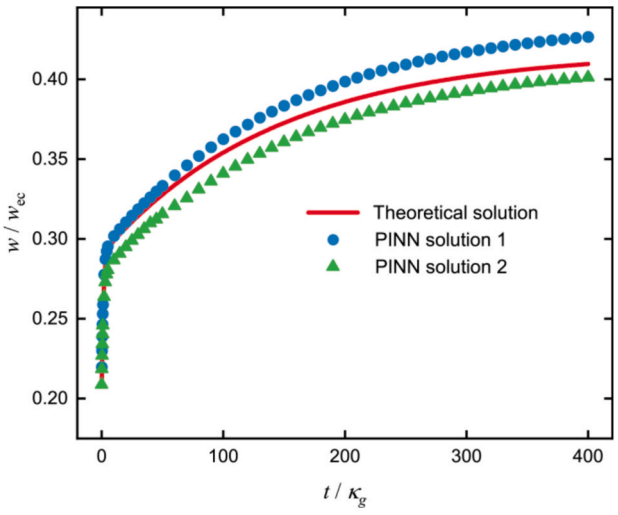

Fig. 4. Comparison of numerical error in the incremental PINNs framework. When the tolerance and the weights of loss items are different, the simulation results of PINNs may have numerical error: PINN solution 1 with the larger time step $\Delta t \geq \kappa_g$, and PINN solution 2 with the larger tolerance (1e-5).

before each loss item). The training process is terminated when a predefined tolerance is met, resulting in a non-zero residual loss. This residual signifies a slight discrepancy in satisfying all objectives simultaneously and is the fundamental source of the approximation error in the calculated displacement increment, Δu, at each step. Second, this incremental method to viscoelasticity is the path-dependent nature of the material model. The error in the deformation increment from a given step is not isolated. It is used to update the internal history variables, which encapsulate the memory of the viscoelastic materials. Consequently, the initial state for the subsequent time step is marginally perturbed from the true physical state. This introduces a compounding effect, where the error from previous steps is propagated. When we lowered the termination tolerance for the loss function (1e-6) and give smaller time step at first ($\Delta t = \kappa_g/10$), the magnitude of this error can be strictly controlled, as shown in Fig. 3c-d.

While error accumulation is an inherent characteristic of this PINNs method, its magnitude can be controlled and minimized through several strategic approaches. The selection of a particular strategy often involves a trade-off between computational cost and desired accuracy. The most direct method is to lower the termination tolerance for the loss function at each time step. Training the network to a smaller residual loss directly reduces the magnitude of the per-step error that is propagated. The primary drawback is a significant increase in the computational cost for each incremental step. Besides, the static weights assigned to different components of the loss function (e.g., physics vs. boundary terms) heavily influence the optimization landscape. Implementing dynamic or adaptive weighting schemes, which adjust the weights during the training process, can prevent any single term from dominating and guide the optimizer to a more balanced and accurate solution. Furthermore, employing hybrid optimization strategies, such as using the Adam optimizer for global exploration followed by a quasi-Newton method like L-BFGS for fine-tuning in this study, can often locate deeper minima in the loss landscape.

4.3. Size-dependent deformation of the beam

The dimensionless deflection against the dimensionless gradient parameter obtained from three methods is illustrated in Fig. 5. For both beam types, the deflection decreases monotonically as the length scale increases. This demonstrates the characteristic size-dependent stiffening effect: materials with a more pronounced microstructure exhibit a higher resistance to bending. Moreover, the SGV can describe time-dependent behavior than traditional theories. When the microstructural and macroscale viscoelasticity are different (different κ_g/τ_g in the figure), the deformation can be very different at the same time. The strain gradient strengthening effect and the deformed pattern can also be seen in Fig. 6. The strain gradient viscoelastic theory gives a larger deflection than the traditional gradient elasticity theory at the same position of the beam with the same gradient parameter. The beam stiffness increases as the gradient parameter increases. Besides, it can be found that the finite element and PINNs can capture the beam deformation accurately along the beam.

To evaluate the practical viability of the proposed framework, we compared the computational resources required for the equivalent mixed-FEM simulations against the PINNs simulations. Table 1 summarizes the wall-clock time for the representative benchmark problems. Computations were performed on a standard computer equipped with an AMD Ryzen 7 5700G CPU (3.80 GHz) to ensure a direct comparison. It is observed that the training time for the PINN exceeds the solution time of the finite element solver for a single forward simulation. However, this computational premium offers distinct advantages in the context of gradient theories. Unlike FEM, which requires complex element formulations to handle the higher-order spatial derivatives in strain gradient theory, PINNs satisfy these smoothness requirements naturally through the activation functions and automatic differentiation. Once the network is trained, the inference cost is effectively

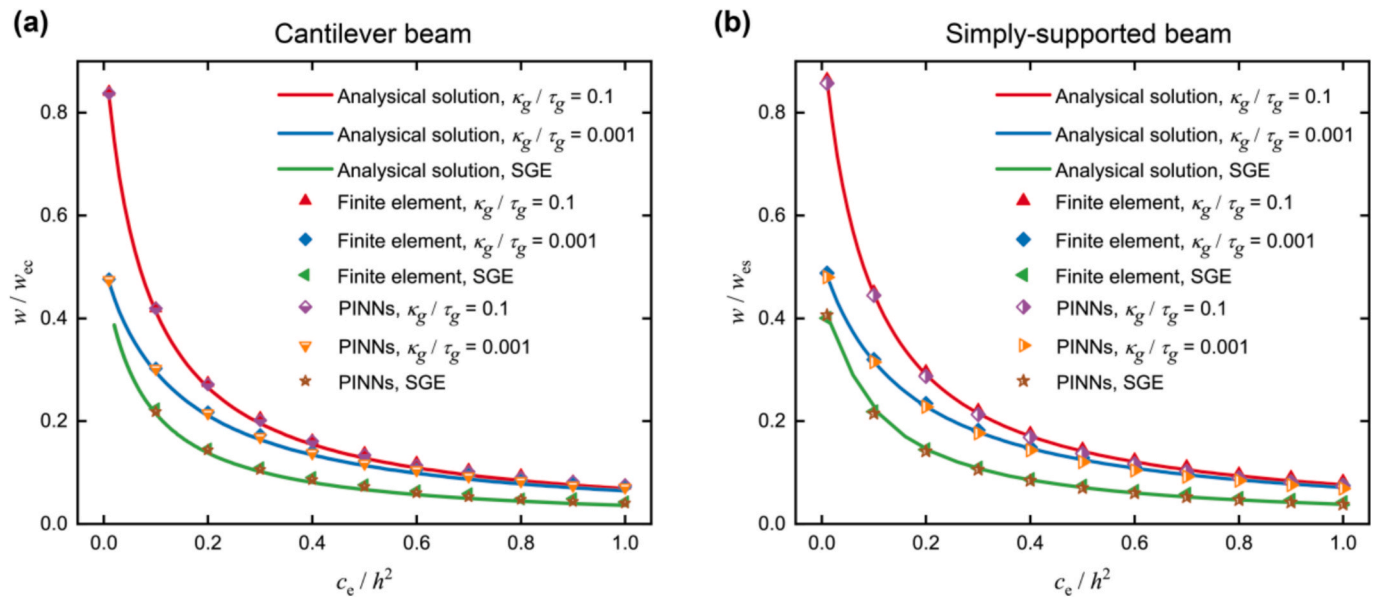

Fig. 5. Dimensionless deflection over the dimensionless gradient parameter for different κ_g/τ_g, with $t = 50\kappa_g$, $E_\infty = E_1$. *SGE* denotes the results of the classical strain gradient elasticity with the elastic modulus $E_0 = E_\infty + E_1$. (a) Deflection at the free end of the cantilever beam. (b) Deflection at the center of the simply-supported beam.

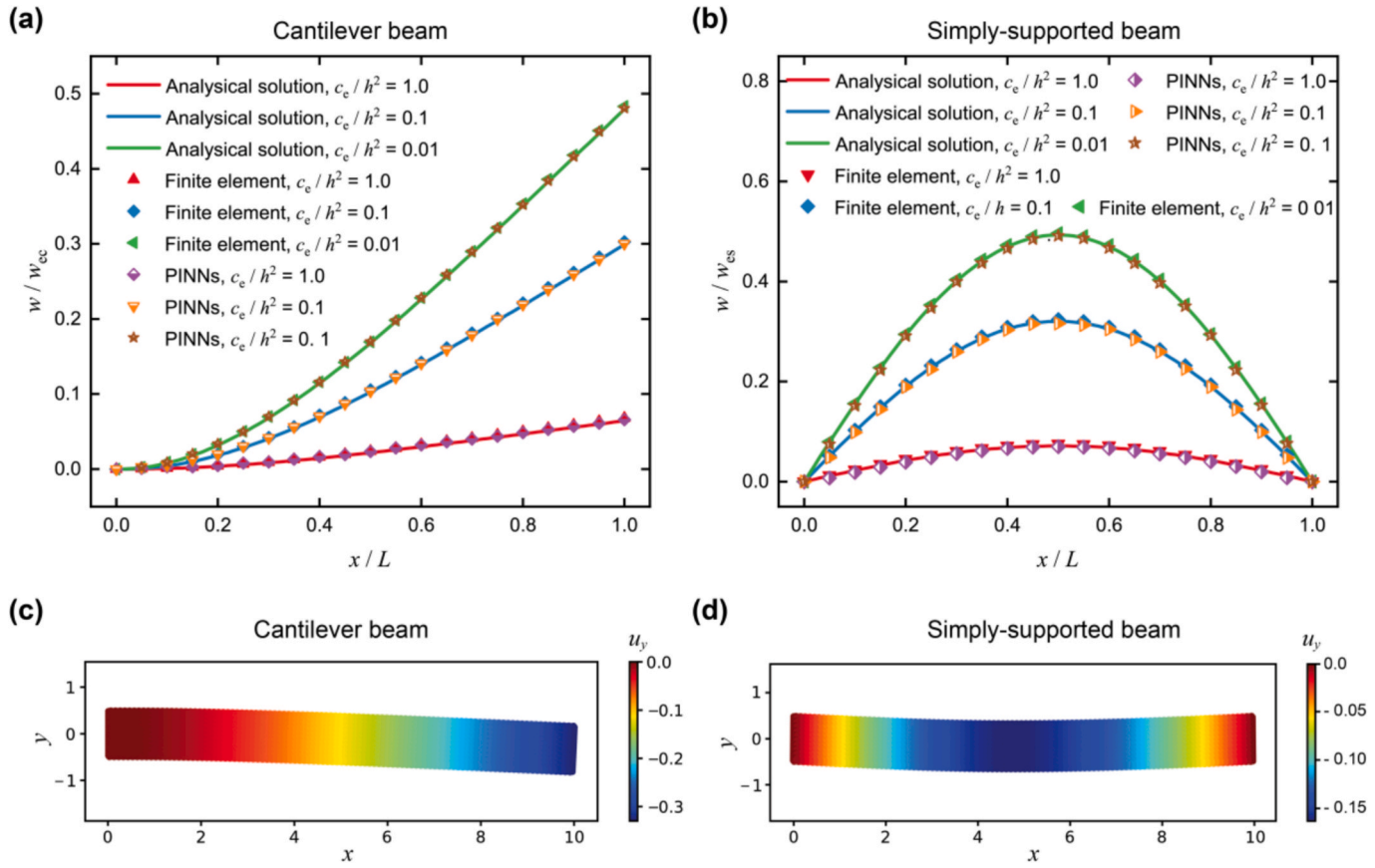

Fig. 6. Dimensionless deflection distribution along the length of beams for different c_e/h^2 at $t = 10\kappa_g$, with $\kappa_g/\tau_g = 0.01$ and $E_\infty = E_1$. (a) The cantilever beam. (b) The simply-supported beam. (c-d) Deformation of beams obtained from PINNs.

instantaneous. Moreover, the trained parameters can serve as an optimized initialization for subsequent simulations with perturbed material properties, which means transfer learning, significantly amortizing the initial training effort across multiple analyses.

4.4. Analysis of coupled spatiotemporal behavior

The power of the SGV framework lies in its ability to capture the coupling between spatial, microstructural effects and temporal evolution. To visualize this, we present spatiotemporal phase portraits in

Table 1
A computational cost comparison of PINN and FEM for beam bending.

Problem	Method	Mesh / Network size	CPU time for time step
Beam bending	FEM	10 × 100 elements	~ 0.3 s
Beam bending	PINN	3 × 40 neurons; 4000 sample points	Train: ≃ 5 mins; Inference: ~ 0.2 s

Fig. 7, which plot the beam deflection as a function of both time and the length scale parameter. These plots reveal a smooth surface where the state of the system (deflection) is uniquely defined by its position in the time–space domain. At t = 0, the response is governed purely by spatial gradients (the strain gradient elasticity limit). As time evolves, the entire surface lifts upwards, representing global viscoelastic relaxation. This temporal evolution, modulated by spatial gradients, manifests as a smooth transition across the phase space, underscoring the natural encoding of spatiotemporal features of SGV. The spatial dimension links microstructural effects to macroscale mechanical responses, while the temporal dimension captures the time-dependent evolution, offering a unified framework that transcends the limitations of classical local theories.

The derived framework offers a useful method to analyze structures with complex microstructural effects, which can be more obvious in advanced structural materials under extreme conditions, such as the extreme high temperatures. SGV can be used to characterize property evolution through the interplay of gradient-viscous coupling. This framework shows promise in characterizing active biological tissues, where hierarchical microstructures induce gradient-like responses coupled with viscoelastic remodeling.

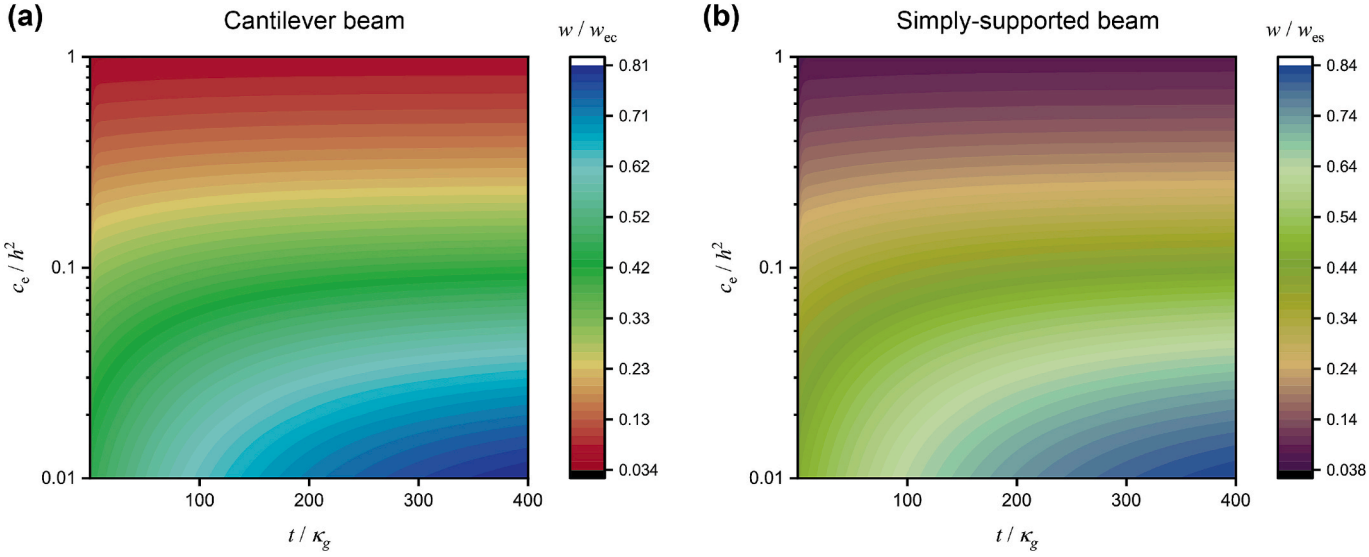

Fig. 7. Spatiotemporal effect captured by the strain gradient viscoelasticity. (a-b) For beams with various microstructures and viscoelasticity.

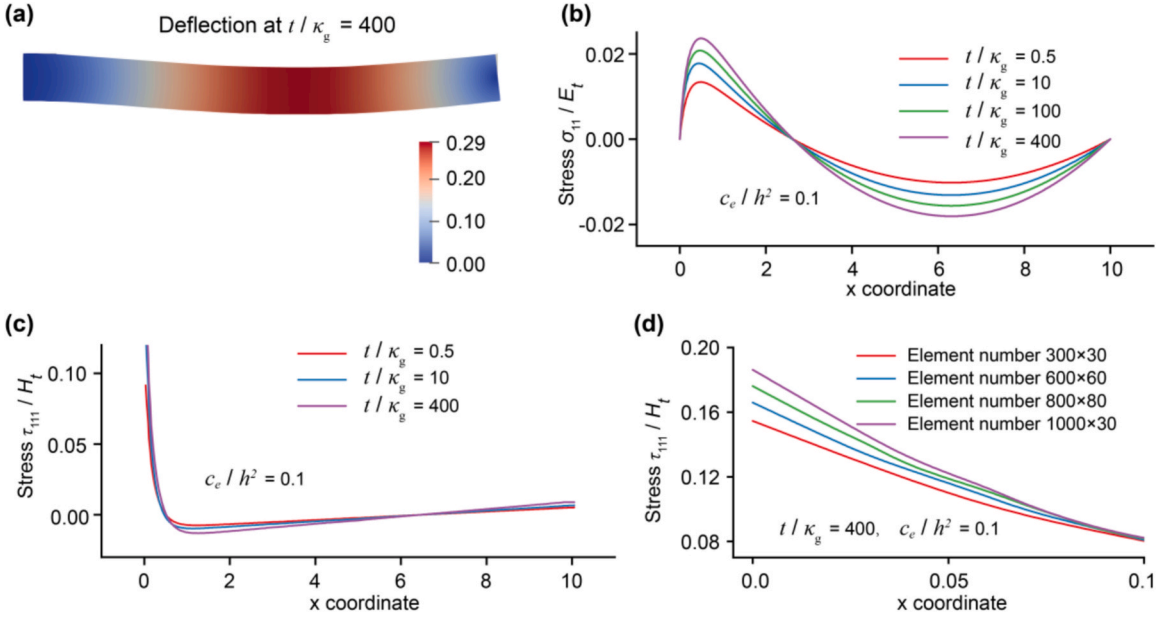

Fig. 8. Bending of a clamped-supported beam, with $\kappa_g/\tau_g = 0.01$, $E_\infty = E_1$, $c_e/h^2 = 0.1$, and the clamped boundary $u_1 = 0$, $u_2 = 0$ and $\phi_{ij} = 0$. (a) Deflection at time $t/\kappa_g = 400$. (b) Dimensionless Cauchy stress σ_{11} along the top surface of the beam. The Cauchy stress is calculated based on the current effective modulus E_t and strain. (c) Dimensionless higher-order stress τ_{111} along the top surface of the beam (excluding results very near the corner). (d) The element size effect on the higher-order result near the clamped corner at time $t/\kappa_g = 400$. The element size is uniform and the size of (a-c) is $L/400$. The higher-order stress is calculated based on the current effective higher-order modulus H_t and strain gradient.

4.5. A statically indeterminate example: The propped cantilever beam

To further evaluate the capability of the proposed mixed-field framework in solving statically indeterminate problems, we consider a propped cantilever beam. The beam is clamped at the left end ($x = 0$) and vertically supported by a roller at the right end ($x = L$). It is subjected to a uniform distributed load. In higher-order continuum theories, the treatment of clamped boundaries presents an interesting discussion regarding the higher-order boundary conditions (Lam et al., 2003; Lazopoulos and Lazopoulos, 2010; Papargyri-Beskou et al., 2003). To demonstrate the critical influence of the micro-boundary conditions on the macroscopic limit, two distinct clamping cases are discussed.

Case A: The micro-clamped boundary, i.e.*, the kinematic micro-constraint.*

In this case, representing an ultra-hard clamp, the material microstructure is assumed to be rigidly gripped by the support. In the 2D finite element framework, this is enforced by fixing both the macroscopic displacements ($u_1 = 0$, $u_2 = 0$) and artificially constraining the nodal strain degrees of freedom ϕ_{ij} to zero at the boundary, which is analogous to $w' = 0$ in 1D theory (Lam et al., 2003; Lazopoulos and Lazopoulos, 2010). The deflection is shown in Fig. 8a. A critical aspect of this benchmark is the stress distribution at the clamped wall. In classical linear elasticity, it is well-established that the fully clamped boundary exhibits a stress singularity at the corners, as the transition from a fixed to a free boundary creates unbounded stresses (Gregory and Gladwell, 1982). In the micro-clamped boundary, the strain gradient framework fundamentally alters this behavior through non-local regularization, as shown in Fig. 8b. The higher-order boundary conditions at the clamped end include the strain degrees of freedom $\phi_{ij} = 0$, and the classical Cauchy normal stress σ_{11} becomes to vanish at the wall. The bending moment at the wall is supported entirely by the higher-order stress τ_{111}, which remains non-zero (Fig. 8c-d). As shown in Fig. 8d, the higher-order stress τ_{111} exhibits a localized singularity at the clamped corner. In the numerical finite element framework, this is verified by the lack of mesh convergence at this specific node: as the element size decreases, the peak value of τ_{111} increases without bound. Consequently, when the classical limit is approached, i.e., $c \to 0$, the strain gradient effect vanishes, leaving the cross-section completely unable to develop a bending moment. This means that the strict micro-clamped assumption is physically meaningful when the strain gradient effect is not neglectable and fails to recover the classical size-independent continuum condition.

Case B: The macro-clamped boundary, i.e.*, the natural micro-constraint.*

In this case, representing a conventional macroscopic wall, the support restricts average translation and rotation but permits the microstructure to deform naturally to carry the load. In the 2D formulation, the standard classical kinematic constraints are applied ($u_1 = 0$, $u_2 = 0$), but the nodal strain degrees of freedom are left unconstrained. This naturally enforces the vanishing of the higher-order work-conjugate tractions, allowing the classical strain components to develop freely. The deflection is shown in Fig. 9a, which exhibits a slightly larger magnitude than the one in Case A. The stress profiles are shown in Fig. 9b-d. By leaving the strain degrees of freedom unconstrained, a non-

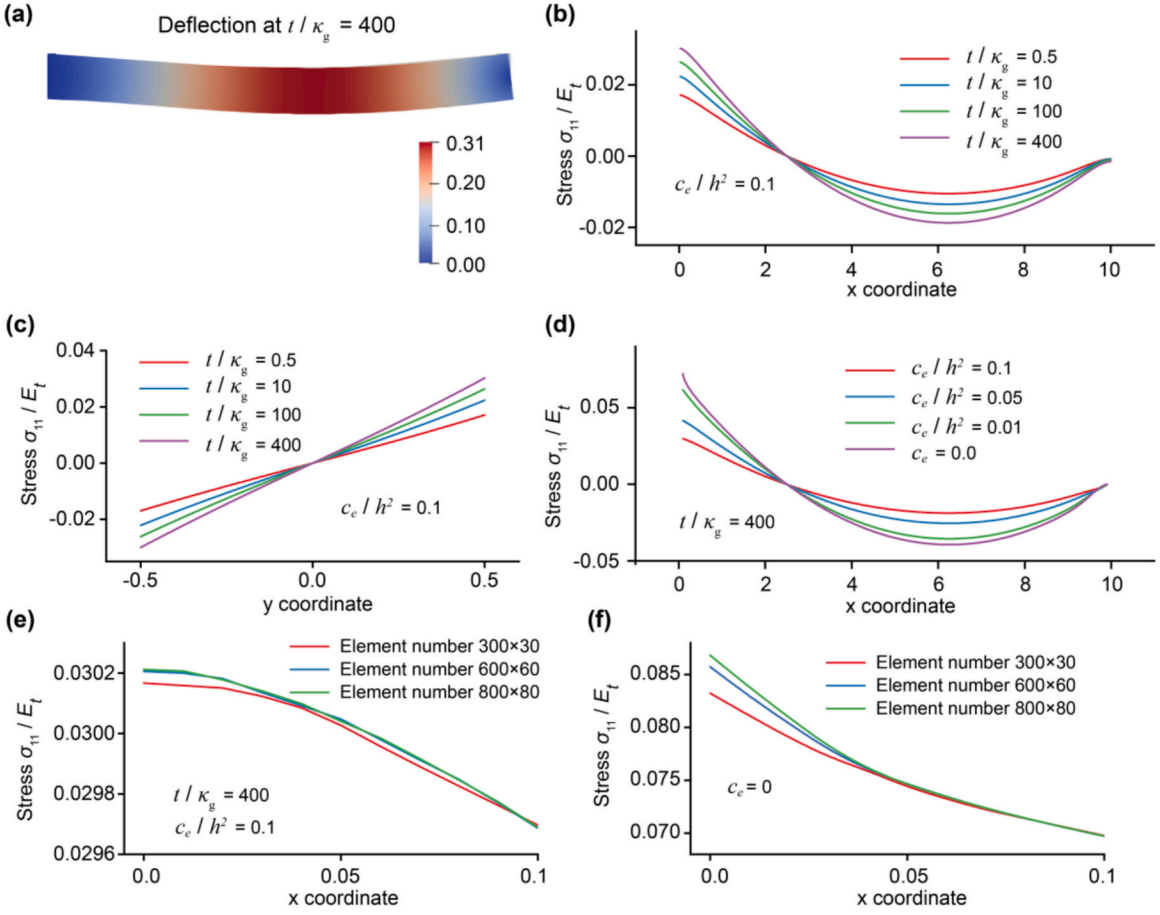

Fig. 9. Bending of a clamped-supported beam, with $\kappa_g/\tau_g = 0.01$, $E_\infty = E_1$, and the clamped boundary $u_1 = 0$, $u_2 = 0$. (a) Deflection at time $t/\kappa_g = 400$. (b) Dimensionless Cauchy stress σ_{11} along the top surface of the beam. (c) Dimensionless Cauchy stress σ_{11} along the clamped boundary at different times. (d) Dimensionless Cauchy stress σ_{11} along the top surface of the beam (excluding results very near the corner) with different microstructural effects. The element size effect on the dimensionless stress near the clamped corner at time $t/\kappa_g = 400$(e) and at the classical elastic condition (f). The element size is uniform and the size of (a-d) is $L/400$.

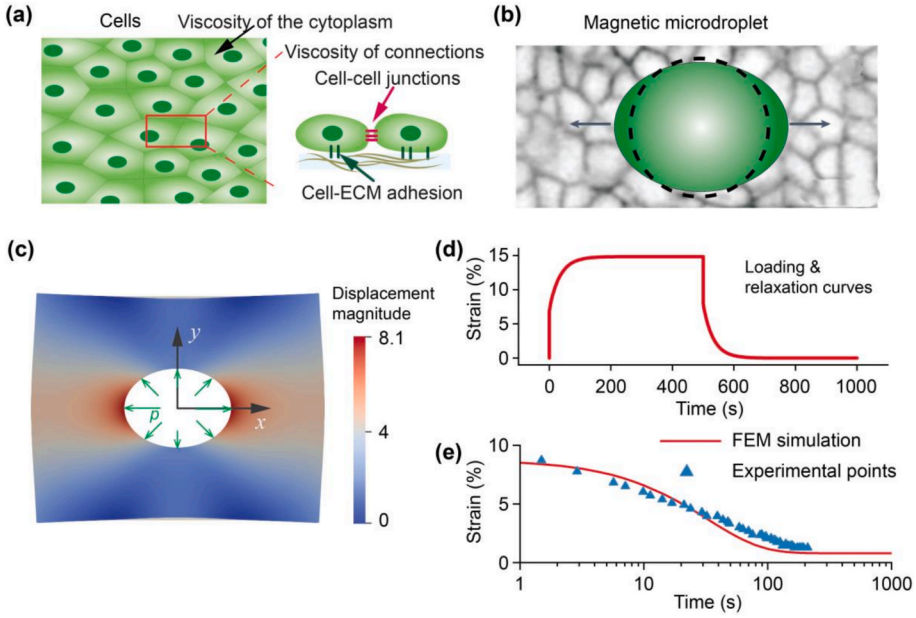

Fig. 10. Spatiotemporal effect of biological tissue and the characterization of strain gradient viscoelasticity. (a) Different viscosities of biological tissue at different scales. There are at least two timescales for viscous characterization: the stress relaxation timescale of cytoplasm is 0.1–1 s, and the timescale of cellular junction is tens of seconds. (b) The magnetic microdroplet used to test the viscoelasticity of biological tissue, where dashed line is the inactive state which is activated in the direction of the arrow. Adapted from (Mongera et al., 2023). (c) Maximum deformed displacement of the rectangular plate with a circle hole, with the pressure around the perimeter boundary of hole $p = p_0[1.0 + 0.8\cos(2\theta)]$ and pressure magnitude $p_0 = E_\infty/10$. The initial geometry of the rectangular plate is a length of $L = 200$ μm and a width of $b = 160$ μm, and the radius of the circular hole is $R = 30$ μm. The symmetric boundary conditions are applied along the symmetry axis ($x = 0$ and $y = 0$). (d) The loading and unloading curve of the problems shown in (c), the strain was the strain magnitude of ε_{11} at the x-direction near the point $[R, 0]$. (e) Comparison of FEM simulation and experimental data (viscoelasticity of yolk, from (Serwane et al., 2017)), which indicates that our strain gradient viscoelasticity can capture the viscoelastic behavior of biological tissue. Time zero corresponds to the start of the relaxation process following the removal of pressure.

vanishing classical stress σ_{11} naturally develops to support the bending moment (Fig. 9c). Moreover, as the material length scale vanishes, the σ_{11} converges to the classical, size-independent continuum solution (Fig. 9d). Furthermore, the 2D continuum model using the strain gradient theory removes the singularities at the extreme clamped corner of the clamped cross-section, yielding a finite, physically bounded stress concentration (Fig. 9e). When the strain gradient effect is neglected (c_e=0), the stress singularity occurs (Fig. 9f), which is the result of a conventional 2D linear elasticity.

5. A potential application: Multiscale viscoelasticity of biological tissues

To demonstrate the capacity of the SGV model to capture the complex, cross-scale mechanical behavior of real physical systems, we apply our PINN framework to a problem inspired by the viscoelasticity of biological tissues. Cellular and tissue mechanics are inherently multiscale, where the viscous response is governed by mechanisms acting at different characteristic length and time scales ((Lin et al., 2025b), Fig. 10a). For instance, the stress relaxation timescale of cytoplasm is quick (0.1–1 s), governed by protein diffusion and network reorganization, while the relaxation of cell–cell junctions and the extracellular matrix (ECM)-cell adhesion is much slower (tens of seconds) (Cheng et al., 2025; Serwane et al., 2017). Classical viscoelasticity, defined by a single constitutive law, cannot fully capture this scale-dependent viscosity. The SGV framework provides a natural solution because its two independent viscosity terms, i.e., the classical viscosity and the higher-order viscosity which linked to microstructure.

A rectangular domain with a circular hole is used to simulate the experiments shown in Fig. 10b (Mongera et al., 2023; Serwane et al., 2017). The rectangular domain with a circular hole is subjected to an internal pressure around the perimeter of the hole, followed by a relaxation process. The mechanical parameters are set to reflect the multi-scale nature of the tissue according to the experiments (Serwane et al., 2017): the moduli $E_\infty = 60$ Pa, $E_1 = 70$ Pa, and viscous parameter $\eta_1 = 1e3$ Pa·s. Other parameters are set as the initial gradient parameter $c_0 = 10$ μm^2, and the microstructural viscous parameter $\zeta_1 = 10^{-2}\eta_1$. The internal pressure is applied instantaneously, where the steady deformation after applied the tension is illustrated in Fig. 10c. Then, the pressure is removed and the subsequent relaxation of the overall strain is monitored, as shown in Fig. 10d. Then, Fig. 10e presents the comparison between the finite element method (FEM) simulation results using the SGV model and representative experimental data for biological tissue (yolk) recovery process after the force acted on the microdroplet is removed. The SGV model can capture the complex profile observed experimentally: it shows a rapid initial decay, accurately matching the sharp strain drop in the experimental data. This is attributed to instant elastic release and the higher-order viscosity term relaxing quickly due to its small relaxation time. Then, the curve transitions to a much slower decay, eventually reaching a steady-state value. This segment is governed by the classical viscosity term with its large relaxation time.

The agreement (Fig. 10e) shows that the SGV framework is possible characterizes the scale-dependent viscosity of biological tissues (Hadzipasic et al., 2024; Mongera et al., 2023). By allowing different length and time scales to evolve independently but coupled within the same continuum model, SGV may provide a powerful tool for predictive modeling in biomechanics and materials science. However, the current SGV framework is formulated within the context of linear kinematics and linear viscoelasticity. While this is sufficient for modeling small-deformation relaxation phenomena, biological tissues and soft polymers often undergo large deformations and active structural remodeling. Future extensions of this work will generalize the theory to finite strain gradient viscoelasticity and incorporate active remodeling

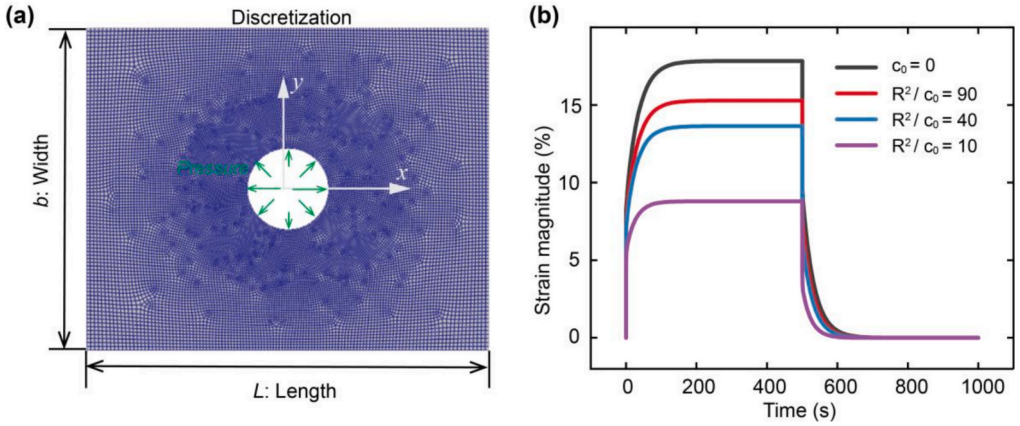

Fig. 11. Size effect capture by SGV. (a) Geometry and discretization of the rectangular plate with a circula hole, with geometry $b/L = 0.8$ and the radius of the hole $R = 0.1L$. (b) Size-dependent mechanical behavior captured by SGV model.

evolution laws, utilizing the thermodynamic and numerical foundations established here.

To differentiate the proposed SGV model from classical size-independent viscoelasticity, we performed a numerical parametric study using the calibrated material parameters. Fig. 11 illustrates the predicted relaxation responses for a rectangular plate with a circular hole, whose geometrical features are: $b/L = 0.8$, the radius of the hole $R = 0.1L$. The discretization for finite element is shown in Fig. 11a, the element size around the hole is $R/15$. The material parameters are the same as previous, i.e., the moduli $E_\infty = 60$ Pa, $E_1 = 70$ Pa, the viscous parameter $\eta_1 = 1e3$ Pa·s, the initial gradient parameter $c_0 = 10\ \mu m^2$, and the microstructural viscous parameter $\zeta_1 = 10^{-2}\eta_1$. The pressure is applied around the perimeter boundary of hole at $t = 0^+$, with $p = p_0[1.0 + 0.8\cos(2\theta)]$ and pressure magnitude $p_0 = E_\infty/10$. This pressure is removed at $t = 500$s. The classical viscoelastic model would predict identical curves regardless of size, as shown by the black curve in Fig. 11b. However, the SGV model predicts a distinct 'stiffer-is-smaller' effect. This demonstrates that the proposed framework captures size-dependent mechanical behavior that classical theories cannot, providing a theoretical basis for future experimental investigations into the micromechanics of tissues.

6. Conclusion

In this study, we presented a comprehensive theoretical and computational study of strain gradient viscoelasticity as a spatiotemporal continuum framework. We have underscored that the power of SGV lies in its intrinsic coupling of a spatial dimension, governed by an internal length scale that models microstructural effects, and a temporal dimension, which captures the hereditary, dissipative evolution of the material response. The primary contributions of this study are threefold. First, we developed two numerical frameworks capable of solving the high-order, integral–differential governing equations of SGV. The mixed-field finite element method provides a rigorous and mature pathway, successfully overcoming the C1 continuity challenge, while the novel variational physics-informed neural network offers a flexible, mesh-free alternative that adeptly handles complex boundary conditions and high-order derivatives. Second, through the benchmark problem of beam bending, we thoroughly validated both numerical methods against the analytical solutions, demonstrating their accuracy and reliability. Our parametric analyses systematically confirmed the ability to predict key physical phenomena: size-dependent stiffening and time-dependent viscoelastic creep. We also introduced the concept of a spatiotemporal phase portrait, providing a clear visualization of how spatial gradient effects and temporal relaxation are fundamentally intertwined. Finally, we demonstrated the physical significance of the SGV model in capturing the complex, multi-scale behavior of soft materials, such as biological tissues. The results highlight the potential of the proposed method to characterize multiscale viscoelastic mechanisms.

The mixed-field formulation increases the computational burden by introducing auxiliary degrees of freedom compared to classical displacement-based elements. However, this is a necessary trade-off to circumvent the stringent C1-continuity requirements of strain gradient theories while using standard mesh generators. In terms of stability, our analyses confirm that the formulation remains well-conditioned for the 2D beam problems considered, allowing for efficient solution via direct sparse solvers without singular behavior. Our comparative analysis indicates that the PINN framework shows promise for problems with higher-order differential and release the mixed-type requirements. The main challenge for the incremental PINN approach remains the management of error propagation, an area ripe for future research. This research opens several avenues for future work. The computational frameworks established here can be extended to three dimensions and to more complex constitutive models incorporating plasticity (strain gradient viscoplasticity) or damage. Furthermore, the demonstrated ability of SGV to model coupled spatiotemporal phenomena makes it a promising tool for application in diverse fields, from predicting the long-term performance of thermal barrier coatings under thermal gradients to modeling the morphogenesis and mechanobiology of living tissues.

CRediT authorship contribution statement

Zhongya Lin: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Conceptualization. **Yanfei Wang:** Writing – review & editing, Visualization, Validation, Software, Funding acquisition, Formal analysis, Conceptualization. **Yueguang Wei:** Writing – review & editing, Validation, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Supports from the National Natural Science Foundation of China (Grant Nos. 12032001, 12432003, and 12472109), and the National Science and Technology Major Project of China (grant no J2022-V-0003-0029) are acknowledged.

Appendix 1

Appendix A. . Solution of the cantilever beam

Considering a rectangular cantilever Bernoulli-Euler beam, and the size relations are $L = 10h$, $h = b$. The fixed end is $x = 0$, and the beam is subjected to a static uniformly distributed lateral load $q(t)$. The deflection $w(x,t)$ of the beam in bending satisfies Eq. (57). The classical boundary conditions are

$$w(0,t) = 0, \quad w'(0,t) = 0, \tag{A1}$$

$$M(L,t) = 0, Q(L,t) = 0. \tag{A2}$$

Using Eq. (56) and $Q = \mathrm{d}M/\mathrm{d}x$, Eq. (A2) also means

$$\begin{aligned} (IE + aE * \mathrm{d}c) * \mathrm{d}w''(L,t) - IE * \mathrm{d}c * \mathrm{d}w^{IV}(L,t) = 0, \\ (IE + aE * \mathrm{d}c) * \mathrm{d}w'''(L,t) - IE * \mathrm{d}c * \mathrm{d}w^{V}(L,t) = 0. \end{aligned} \tag{A3}$$

And the non-classical boundary conditions are taken as

$$\begin{aligned} M_h(L,t) = 0, \text{ that is } w'''(L,t) = 0, \\ w''(0,t) = 0. \end{aligned} \tag{A4}$$

To obtain viscoelastic gradient solutions, the Laplace transform is applied. In the Laplace phase space, the governing equation (the Laplace transformation of Eq. (57)) can be obtained

$$(s\overline{E}I + as\overline{E}s\overline{c})\overline{w}^{\mathrm{IV}} - Is\overline{E}s\overline{c}\overline{w}^{\mathrm{VI}} = \overline{q}. \tag{A5}$$

The Laplace transformed classical boundary conditions, Eqs. (A1)-(A3), are

$$\begin{aligned} \overline{w}(0,s) = 0, \\ \overline{w'}(0,s) = 0, \\ s(I\overline{E} + a\overline{E}s\overline{c})\overline{w''}(L,s) - Is\overline{E}s\overline{c}\overline{w}^{IV}(L,s) = 0, \\ s(I\overline{E} + a\overline{E}s\overline{c})\overline{w'''}(L,s) - Is\overline{E}s\overline{c}\overline{w}^{V}(L,s) = 0. \end{aligned} \tag{A6}$$

And the non-classical boundary conditions

$$\begin{aligned} \overline{w'''}(L,s) = 0, \\ \overline{w''}(0,s) = 0. \end{aligned} \tag{A7}$$

The general solution of differential equations is the sum of the solution of its homogeneous part, i.e., the one with $q(t) = 0$, and a particular solution. Setting the former part of the solution of Eq. (A5) as

$$\overline{w}_h(x) = \overline{a}_1 + \overline{a}_2 x + \overline{a}_3 x^2 + \overline{a}_4 x^3 + \overline{a}_5 \exp\left(\sqrt{(I + as\overline{c})/Is\overline{c}}\, x\right) + \overline{a}_6 \exp\left(-\sqrt{(I + as\overline{c})/Is\overline{c}}\, x\right). \tag{A8}$$

where $\overline{a}_i$, $i = 1, \ldots, 6$ are parameters to be determined, and they are the function of time. And a particular solution is taken as

$$\overline{w}_{\mathrm{p}}(x) = \frac{\overline{q}}{24(s\overline{E}I + as\overline{E}s\overline{c})} x^4. \tag{A9}$$

Thus, solutions of parameters $\overline{a}_i$ can be obtained by applying the boundary conditions Eqs. (A6) and (A7).

$$\begin{aligned} a_1 &= \frac{h\left(2h + gL^2 + mL^2\right)}{2(g+m)^3}\overline{q}, \\ a_2 &= -\frac{\left(e^{2kL} - 1\right)\left(2h + gL^2 + mL^2\right)}{2k(g+m)^2(e^{2kL} + 1)}\overline{q}, \\ a_3 &= \frac{\left(2h + gL^2 + mL^2\right)}{4(g+m)^2}\overline{q}, \\ a_4 &= -\frac{L}{6(g+m)}\overline{q}, \\ a_5 &= -\frac{2e^{-kL}k^5\left(6/k^2 + 3L^2\right)}{-e^{kL}k^5\left(12L(g+m)e^{-kL}k^3 - 12hk^4\right) - e^{-kL}k^5\left(-12L(g+m)e^{kL}k^3 - 12hk^4\right)}\overline{q}, \\ a_6 &= -\frac{\mathrm{e}^{2kL}h\left(2h + gL^2 + mL^2\right)}{2(1 + \mathrm{e}^{2kL})(g+m)^3}\overline{q}. \end{aligned} \tag{A10}$$

where $g = s\overline{E}I$, $h = s\overline{A}I = s^2\overline{c}\overline{E}I$, $m = s\overline{A}a = s^2\overline{c}\overline{E}a$, and $k = \sqrt{(g+m)/h}$. Then, using the standard three-parameter viscoelastic models as shown in Fig. 1, and the relations $H_\infty = c_e E_\infty$, $H_1 = c_e E_1$, $\overline{E}$ and $\overline{c}$ are written as

$$\overline{E} = \frac{E_\infty}{s} + \frac{E_1\tau_g}{1+\tau_g s}, \text{(A11)}.$$

$$\overline{c} = c_e\left(\frac{E_\infty}{s} + \frac{E_1\kappa_g}{1+\kappa_g s}\right)/s\left(\frac{E_\infty}{s} + \frac{E_1\tau_g}{1+\tau_g s}\right), \text{(A12)}.$$

where τ_g is the classical relaxation time, κ_g is the relaxation time of higher-order viscoelastic models, c_e is an elastic gradient parameter which links the classical moduli and higher-order moduli. The gradient parameter c can be obtained.

$$c(t) = c_e\left\{1 + \frac{E_1(\kappa_g-\tau_g)}{E_\infty\kappa_g-(E_\infty+E_1)\tau_g}\left[e^{\frac{-t}{\kappa_g}} - e^{\frac{-E_\infty t}{(E_\infty+E_1)\tau_g}}\right]\right\}. \text{(A13)}.$$

In view of Eqs. (A8)-(A12), the general solution $\overline{w} = \overline{w}_h + \overline{w}_p$ of deflection in the Laplace phase space can be obtained. Then, applying the inverse Laplace transformation, the general solution of beam bending deformation can be obtained. It should be mentioned that the deflection is obtained using a numerical inverse Laplace transformation method in a MATLAB program.

Appendix B. . Solution of the simply-supported beam

This section deals with a simply-supported beam of length L subjected to a uniformly distributed lateral load $q(t)$. The classical boundary conditions are

$$w(0,t) = 0, \quad w(L,t) = 0, \tag{B1}$$

$$M(0,t) = 0, M(L,t) = 0. \tag{B2}$$

Eq. (56), Eq. (B2) also means

$$\begin{aligned}(IE + aE * \mathrm{d}c) * \mathrm{d}w''(0,t) - IE * \mathrm{d}c * \mathrm{d}w^{IV}(0,t) = 0,\\ (IE + aE * \mathrm{d}c) * \mathrm{d}w''(L,t) - IE * \mathrm{d}c * \mathrm{d}w^{IV}(L,t) = 0,\end{aligned} \tag{B3}$$

The non-classical boundary conditions are considered as

$$\begin{aligned}w''(0,t) = 0,\\ w''(L,t) = 0.\end{aligned} \tag{B4}$$

Thus, the Laplace transformed classical boundary conditions, Eqs. (B1)-(B3), are

$$\begin{aligned}\overline{w}(0,s) = 0,\\ \overline{w}(L,s) = 0,\\ s(I\overline{E} + a\overline{E}s\overline{c})\overline{w''}(0,s) - Is\overline{E}s\overline{c}\overline{w}^{IV}(0,s) = 0,\\ s(I\overline{E} + a\overline{E}s\overline{c})\overline{w''}(L,s) - Is\overline{E}s\overline{c}\overline{w}^{IV}(L,s) = 0.\end{aligned} \tag{B5}$$

The Laplace transformed non-classical boundary conditions

$$\begin{aligned}\overline{w}''(0,s) = 0,\\ \overline{w}''(L,s) = 0.\end{aligned} \tag{B6}$$

The general solution of the homogeneous part and a particular solution of differential equations are still taken as Eq. (A8) and Eq. (A9), respectively. Hence, the solutions of $\overline{a}_i, i = 1, \ldots, 6$ can be obtained using the boundary conditions Eqs. (B5)-(B6)

$$\begin{aligned}a_1 &= \frac{h^2}{(g+m)^3}\overline{q},\\ a_2 &= \frac{L(-12h + gL^2 + mL^2)}{24(g+m)^2}\overline{q},\\ a_3 &= \frac{h}{2(g+m)^2}\overline{q},\\ a_4 &= -\frac{L}{12(g+m)}\overline{q},\\ a_5 &= -\frac{h^2}{(1+e^{kL})(g+m)^3}\overline{q},\\ a_6 &= -\frac{e^{kL}h^2}{(1+e^{2kL})(g+m)^3}\overline{q},\end{aligned} \tag{B7}$$

where $g = s\overline{E}I$, $h = s\overline{A}I = s^2\overline{c}\overline{E}I$, $m = s\overline{A}a = s^2\overline{c}\overline{E}a$, and $k = \sqrt{(g+m)/h}$. $\overline{E}$ and $\overline{c}$ are written as Eqs. (A11) and (A12), respectively.

Then, the general solution of beam bending deformation can be obtained by using the numerical inverse Laplace transformation with the help of a MATLAB program. Furthermore, in the previous study (Lin and Wei, 2020), a correspondence principle between the strain gradient viscoelasticity and the classical strain gradient elasticity is summarized. Therefore, the strain gradient viscoelastic solutions of the cantilever beam bending can also be obtained directly using the correspondence principle. That is, the strain gradient viscoelastic solution in Laplace phase space can be obtained by replacing E, c in classical strain gradient elastic solution Eq.(61) with $s\bar{c}(s)$ and $s\bar{E}(s)$, respectively.

Data availability

No data was used for the research described in the article.

References

Abueidda, D.W., Koric, S., Al-Rub, R.A., Parrott, C.M., James, K.A., Sobh, N.A., 2022. A deep learning energy method for hyperelasticity and viscoelasticity. Eur. J. Mech. A Solids 95, 104639.

Alnæs, M.S., Logg, A., Ølgaard, K.B., Rognes, M.E., Wells, G.N., 2014. Unified form language. ACM Trans. Math. Softw. 40, 1–37.

Amanatidou, E., Aravas, N., 2002. Mixed finite element formulations of strain-gradient elasticity problems. Comput. Methods Appl. Mech. Eng. 191, 1723–1751.

Askes, H., Aifantis, E.C., 2011. Gradient elasticity in statics and dynamics: an overview of formulations, length scale identification procedures, finite element implementations and new results. Int. J. Solids Struct. 48, 1962–1990.

Askes, H., Morata, I., Aifantis, E.C., 2008. Finite element analysis with staggered gradient elasticity. Comput. Struct. 86, 1266–1279.

Bai, J., Lin, Z., Wang, Y., Wen, J., Liu, Y., Rabczuk, T., Gu, Y., Feng, X.-Q., 2025. Energy-based physics-informed neural network for frictionless contact problems under large deformation. Comput. Methods Appl. Mech. Eng. 437, 117787.

Barrett, J.W., Elliott, C.M., 1986. Finite Element Approximation of the Dirichlet Problem using the Boundary Penalty Method. Numer. Math. 49, 343–366.

Berrone, S., Canuto, C., Pintore, M., Sukumar, N., 2023. Enforcing Dirichlet Boundary Conditions in Physics-Informed Neural Networks and Variational Physics-Informed Neural Networks. Heliyon 9, e18820.

Chen, S.H., Wang, T.C., 2002. A new deformation theory with strain gradient effects. Int. J. Plast 18, 971–995.

Cheng, B., Li, M., Lin, M., Guo, H., Xu, F., 2025. Mechanobiology across Timescales. Nat. Rev. Phys. 1–24.

Christensen, R.M., 1982. Theory of viscoelasticity, 2nd ed. Dover Publications, Mineola N.Y., xii, p. 364.

Chen, D., Wang, Y., Zhou, H., Huang, Z., Zhang, Y., Guo, C.F., Zhou, H., 2022. Current and Future Trends for Polymer Micro/Nanoprocessing in Industrial applications. Adv. Mater. 34, 2200903.

Ding, K., Lin, Z., Wei, Y., 2023. The strain gradient viscoelasticity full field solution of mode-III crack problem. Int. J. Fract. 242, 71–83.

Ebrahimi, F., Barati, M.R., 2017. Hygrothermal effects on vibration characteristics of viscoelastic FG nanobeams based on nonlocal strain gradient theory. Compos. Struct. 159, 433–444.

Gao, X.L., Park, S.K., 2007. Variational formulation of a simplified strain gradient elasticity theory and its application to a pressurized thick-walled cylinder problem. Int. J. Solids Struct. 44, 7486–7499.

Gregory, R.D., Gladwell, I., 1982. The cantilever beam under tension, bending or flexure at infinity. J Elasticity 12, 317–343.

Grossmann, T.G., Komorowska, U.J., Latz, J., Schönlieb, C.-B., 2024. Can Physics-Informed Neural Networks beat the Finite Element Method? J. Micromech. Microeng. 89, 143–174.

Gusev, A.A., Bernhard, T., 2024. Molecular Model for Linear Viscoelastic Properties of Entangled Polymer Networks. Macromolecules 57, 10152–10163.

Hadzipasic, M., Zhang, S., Huang, Z., Passaro, R., Sten, M.S., Shankar, G.M., Nia, H.T., 2024. Emergence of nanoscale viscoelasticity from single cancer cells to established tumors. Biomaterials 305, 122431.

Huang, C., Yang, B., Peng, X., Chen, S., 2020. Plastic Deformation and Hardening Mechanisms of a Nano-twinned Cubic Boron Nitride Ceramic. ACS Appl. Mater. Interfaces 12, 50161–50175.

Huang, G., Tang, Z., Peng, S., Zhang, P., Sun, T., Wei, W., Zeng, L., Guo, H., Guo, H., Meng, G., 2022. Modification of Hydrophobic Hydrogels into a strongly Adhesive and tough Hydrogel by Electrostatic Interaction. Macromolecules 55, 156–165.

Jiang, Y., Li, L., Hu, Y., 2023. Strain gradient viscoelasticity theory of polymer networks. Int. J. Eng. Sci. 192, 103937.

Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., Yang, L., 2021. Physics-Informed Machine Learning. Nat Rev Phys 3, 422–440.

Baratta, I.A., Dean, J.P., Dokken, J.S., Habera, M., Hale, J.S., Richardson, C.N., Rognes, M.E., Scroggs, M.W., Sime, N., Wells, G.N., 2025. DOLFINx: The next generation FEniCS problem solving environment. Zenodo.

Kingma, D.P., Ba, J., 2017. Adam: A Method for Stochastic Optimization. https://doi.org/10.48550/arXiv.1412.6980.

Lam, D.C.C., Yang, F., Chong, A.C.M., Wang, J., Tong, P., 2003. Experiments and theory in strain gradient elasticity. J. Mech. Phys. Solids 51, 1477–1508.

Lazar, M., Maugin, G.A., Aifantis, E.C., 2006. Dislocations in second strain gradient elasticity. Int. J. Solids Struct. 43, 1787–1817.

Lazopoulos, K.A., Lazopoulos, A.K., 2010. Bending and buckling of thin strain gradient elastic beams. Eur. J. Mech. A. Solids 29, 837–843.

Li, H., Miao, Y., Khodaei, Z.S., Aliabadi, M.H., 2024. Finite-PINN: A Physics-Informed Neural Network Architecture for Solving Solid Mechanics Problems with General Geometries. https://doi.org/10.48550/arXiv.2412.09453.

Lin, Z., Bai, J., Li, S., Chen, X., Li, B., Feng, X.-Q., 2026. A Physics-Informed Neural Network Framework for Simulating Creep Buckling in growing Viscoelastic Biological Tissues. Comput. Methods Appl. Mech. Eng. 452, 118715.

Lin, Z., Ding, K., Ma, H., Wei, Y., 2025a. A Trans-scale Shear-lag Model for Characterizing the size effect and Viscoelasticity of Staggered Shells. Acta Mech. Solida Sin. 1–15.

Lin, Z., Huang, W., Li, S., Wang, M., Bai, J., Chen, X., Feng, X.-Q., 2025b. Mechanobiological Modeling of Viscoelasticity in Soft Tissue Growth and Morphogenesis. J. Mech. Phys. Solids 196, 106032.

Lin, Z., Wei, Y., 2020. A strain gradient linear viscoelasticity theory. Int. J. Solids Struct. 203, 197–209.

Liu, H.Y., Wei, Y.G., Liang, L.H., Liu, X.H., Wang, Y.B., Ma, H.S., 2018. Damage characterization model of ceramic coating systems based on energy analysis and bending tests. Ceram. Int. 44, 4807–4813.

Lurie, S., Solyaev, Y., 2018. Revisiting bending theories of elastic gradient beams. Int. J. Eng. Sci. 126, 1–21.

Scroggs, M.W., Baratta, I.A., Richardson, C.N., Wells, G.N., 2022a. Basix: a runtime finite element basis evaluation library. Journal of Open Source Software 7, 3982.

Mindlin, R.D., 1964. Micro-structure in linear elasticity. Arch. Ration. Mech. Anal. 16, 51–78.

Mindlin, R.D., 1965. Second gradient of strain and surface-tension in linear elasticity. Int. J. Solids Struct. 1, 417–438.

Mindlin, R.D., Eshel, N.N., 1968. On first strain-gradient theories in linear elasticity. Int. J. Solids Struct. 4, 109–124.

Mongera, A., Pochitaloff, M., Gustafson, H.J., Stooke-Vaughan, G.A., Rowghanian, P., Kim, S., Campàs, O., 2023. Mechanics of the cellular microenvironment as probed by cells in vivo during zebrafish presomitic mesoderm differentiation. Nat. Mater. 22, 135–143.

Nguyen-Thanh, V.M., Zhuang, X., Rabczuk, T., 2020. A Deep Energy Method for Finite Deformation Hyperelasticity. Eur. J. Mech. A. Solids 80, 103874.

Papargyri-Beskou, S., Tsepoura, K.G., Polyzos, D., Beskos, D.E., 2003. Bending and stability analysis of gradient elastic beams. Int. J. Solids Struct. 40, 385–400.

Polizzotto, C., 2016. A note on the higher order strain and stress tensors within deformation gradient elasticity theories: Physical interpretations and comparisons. Int. J. Solids Struct. 90, 116–121.

Polyzos, D., Fotiadis, D.I., 2012. Derivation of Mindlin's first and second strain gradient elastic theory via simple lattice and continuum models. Int. J. Solids Struct. 49, 470–480.

Rahali, Y., Reda, H., Vieille, B., Lakiss, H., Ganghoffer, J.-F., 2021. Second Gradient Linear and Nonlinear Constitutive Models of Architectured Materials: Static and Dynamic Behaviors. Mathematical Applications in Continuum and Structural Mechanics 53–71.

Rahali, Y., Reis, F.D., Ganghoffer, J.-F., 2017. Multiscale homogenization schemes for the construction of second-order grade anisotropic continuum media of architectured materials. Int. J. Multiscale Comput. Eng. 15.

Reddy, J.N., 2007. Nonlocal theories for bending, buckling and vibration of beams. Int. J. Eng. Sci. 45, 288–307.

Samaniego, E., Anitescu, C., Goswami, S., Nguyen-Thanh, V.M., Guo, H., Hamdia, K., Zhuang, X., Rabczuk, T., 2020. An Energy Approach to the solution of Partial Differential Equations in Computational Mechanics via Machine Learning: Concepts, Implementation and applications. Comput. Methods Appl. Mech. Eng. 362, 112790.

Scroggs, M.W., Dokken, J.S., Richardson, C.N., Wells, G.N., 2022b. Construction of Arbitrary Order Finite Element Degree-of-Freedom Maps on Polygonal and Polyhedral Cell Meshes. ACM Trans. Math. Softw. 48, 1–23.

An, S., Joshi, B., Yarin, A.L., Swihart, M.T., Yoon, S.S., 2020. Supersonic Cold Spraying for Energy and Environmental applications: One-step Scalable Coating Technology for Advanced Micro- and Nanotextured Materials. Adv. Mater. 32, 1905028.

Serwane, F., Mongera, A., Rowghanian, P., Kealhofer, D.A., Lucio, A.A., Hockenbery, Z. M., Campàs, O., 2017. In vivo quantification of spatially varying mechanical properties in developing tissues. Nat. Methods 14, 181–186.

Sukumar, N., Srivastava, A., 2022. Exact imposition of boundary conditions with distance functions in physics-informed deep neural networks. Comput. Methods Appl. Mech. Eng. 389, 114333.

Wang, Y., Bai, J., Lin, Z., Wang, Q., Anitescu, C., Sun, J., Eshaghi, M.S., Gu, Y., Feng, X.-Q., Zhuang, X., Rabczuk, T., Liu, Y., 2024. Artificial Intelligence for Partial

Differential Equations in Computational Mechanics: A Review. https://doi.org/10.48550/arXiv.2410.19843.
Wang, Y., Wei, Y., 2024. Strain energy density maximization principle for material design and the reflection in trans-scale continuum theory. J. Mech. Phys. Solids 193, 105912.
Wei, Y., 2006. A new finite element method for strain gradient theories and applications to fracture analyses. Eur. J. Mech. A. Solids 25, 897–913.
Wei, Y.G., Hutchinson, J.W., 1997. Steady-state crack growth and work of fracture for solids characterized by strain gradient plasticity. J. Mech. Phys. Solids 45, 1253–1273.
Yu, Z., Lin, Z., Wei, Y., 2021. Closed-form functions of cross-scale indentation scaling relationships based on a strain gradient plasticity theory. Phil. Mag. 79, 1–22.
Zhu, F., Song, S., Reddy, K.M., Hirata, A., Chen, M., 2018. Spatial heterogeneity as the structure feature for structure–property relationship of metallic glasses. Nat. Commun. 9, 1–7.
Zhuang, X., Guo, H., Alajlan, N., Zhu, H., Rabczuk, T., 2021. Deep Autoencoder based Energy Method for the Bending, Vibration, and Buckling Analysis of Kirchhoff Plates with transfer Learning. Eur. J. Mech. A. Solids 87, 104225.
Zybell, L., Mühlich, U., Kuna, M., Zhang, Z.L., 2012. A three-dimensional finite element for gradient elasticity based on a mixed-type formulation. Comput. Mater. Sci 52, 268–273.