Modeling the interaction between instabilities and functional degradation in shape memory alloys

Mohsen Rezaee-Hajidehi[a,*], Maciej Ryś[b,a]

[a]Institute of Fundamental Technological Research (IPPT), Polish Academy of Sciences, Pawińskiego 5B, 02-106 Warsaw, Poland.
[b]NOMATEN Centre of Excellence, National Center for Nuclear Research (NCBJ), A. Sołtana 7, Świerk/Otwock, 05-400, Poland.

Abstract

Localization of the stress-induced martensitic phase transformation plays an important role in the fatigue behavior of shape memory alloys (SMAs). The phenomenon of return-point memory that is observed during the subloop deformation of a partially-transformed SMA is a clear manifestation of the interaction between localized phase transformation and degradation of the functional properties. The present study aims to demonstrate this structure–material interaction in the modeling of return-point memory. It seems that this crucial aspect has been overlooked in previous modeling studies. For this purpose, we developed a gradient-enhanced model of pseudoelasticity that incorporates the degradation of functional properties in its constitutive description. The model is employed to reproduce the hierarchical return-point memory in a pseudoelastic NiTi wire under isothermal uniaxial tension with nested subloops. Additionally, a detailed analysis is carried out for a NiTi strip with more complex transformation pattern. Our study highlights the subtle morphological changes of phase transformation under different loading scenarios and the resulting implications for return-point memory.

Keywords: Shape memory alloys; Phase transformation; Functional degradation; Propagating instabilities; Subloop deformation; Modeling

1. Introduction

The practical interest in shape memory alloys (SMAs), especially NiTi, stems from their ability to withstand and recover large strains. This ability is exhibited through mechanical loading and unloading at sufficiently high temperatures (pseudoelasticity) or through mechanical loading and unloading followed by heating (shape memory effect). The underlying mechanism is the crystallographically reversible martensitic phase transformation that occurs between the austenitic parent phase (stable at higher temperatures, possessing higher crystal symmetry) and the martensitic product phase (stable at lower temperatures, possessing lower crystal symmetry) [1]. By leveraging the unique characteristics of SMAs, they have found a broad range of applications across various fields, from micro-scale biomedical devices to macro-scale industrial components [2, 3]. The operational lifespan of SMAs in most of the applications involves enduring cyclic mechanical/thermal loadings,

*Corresponding author.
Email addresses: mrezaee@ippt.pan.pl (Mohsen Rezaee-Hajidehi), maciej.rys@ncbj.gov.pl (Maciej Ryś)

which highlights the great importance of identifying their fatigue behavior. It is well-recognized that, due to the martensitic phase transformation, fatigue in SMAs is more complex than in common engineering metals and is mainly classified into two aspects: degradation of functional properties (such as recoverable strain, transformation stress, and hysteresis loop area), known as functional fatigue, and the evolution of damage in the material, known as structural fatigue [4]. This complex nature demands special attention and, as a result, has prompted a tremendous number of studies that focus on the fatigue characterization of SMAs from a variety of perspectives and on the underlying micromechanical processes [3–13].

Stress-induced martensitic transformation in pseudoelastic NiTi appears (typically, in tension-dominated loadings) as localized instabilities in the form of martensite bands, and subsequently progresses via propagation of the instabilities in the form of patterned interfaces (macroscopic transformation fronts) that separate the domains of low-strained austenite and high-strained martensite, e.g., [14–17]. Due to the high strain incompatibilities that exist within the transformation front and the ensuing large local stresses, it can be reasonably inferred that propagating instabilities can vitally influence both the functional fatigue and structural fatigue of the material. Despite the longstanding recognition of this crucial aspect [4, 18–21], its direct validation was provided only a few years ago in the experiments conducted by Zheng et al. [8, 22, 23]. It was demonstrated that in view of the repetitive nucleation and propagation of the localized transformation in NiTi strips under cyclic uniaxial tension, a rapid degradation of pseudoelasticity occurs that accelerates the fatigue crack initiation and fatigue failure.

An interesting manifestation of the interaction between propagating instabilities and functional degradation is found in the subloop deformation behavior of a partially-transformed SMA specimen under displacement-controlled loading. The subloop behavior has been extensively investigated experimentally, notably for NiTi [18, 19, 23–26] but also for other SMAs [27, 28]. Fig. 1(a), reproduced from Tobushi et al. [24], depicts the global mechanical response of a NiTi wire subjected to subloop paths. For a more intuitive description of the phenomenon, hypothetical schematics of the corresponding transformation front evolution are provided in Fig. 1(b). As the front propagates along the wire, it leaves behind transformation-induced microstructural defects, such as dislocations and stabilized (locked-in) martensite. During the subloop unloading (for instance, the first subloop, which starts at point A), the front travels backward over an already swept zone (from A to B), hence intensifying the generated defects. Accordingly, during the subloop reloading, the propagation of the front over the twice-swept zone occurs with a lower stress level compared to the original transformation plateau. Upon entering the pristine zone (at point A), which is virtually free of transformation-induced defects, the front experiences the transformation-onset stress characteristic to the initial material state. This causes the stress to catch up with the original plateau by passing through the subloop unloading point. This trait is known as the return-point memory. The process repeats in the subsequent subloops and culminates in an intriguing hierarchical return-point memory.

Motivated by the experimental results, numerous attempts have been made to develop SMA models capable of capturing the phenomenon of return-point memory during the subloop deformation, either through incorporating the permanent strain contribution and degradation of functional properties [29–32] or by merely refining the constitutive equations of (non-cyclic) model of pseudoe-

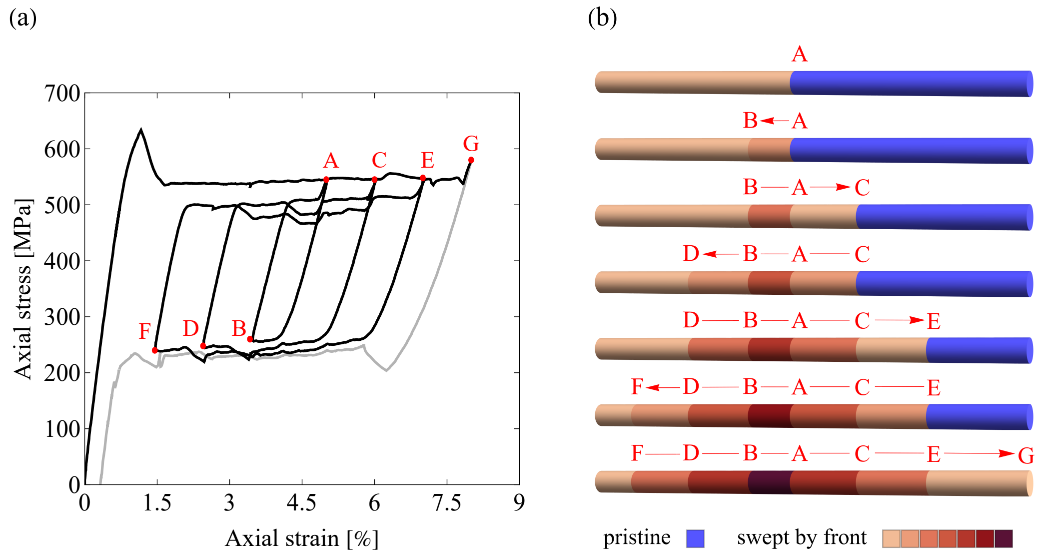


Figure 1: Return-point memory in NiTi wire subjected to uniaxial tension with three nested subloop paths: (a) the structural stress–strain response, and (b) hypothetical schematics of the corresponding transformation front evolution. The stress–strain response in panel (a) is reproduced from Tobushi et al. [24] (courtesy of R. Matsui). The red arrows in panel (b) indicate the trajectory of the front propagation, and the color scales quantify the recurrence of the front's traversal over the wire's segments.

lasticity [25, 33, 34]. In fact, a physically-relevant approach for modeling the return-point memory should hinge on the interaction between the propagating instabilities (structural inhomogeneities) and the functional degradation of the material. Nevertheless, most of the existing models (including those referenced above) postulate a homogeneous martensitic phase transformation, while addressing a problem with a transformation of localized nature. Albeit this simplifies the computations significantly, it is not a plausible assumption in the present context. To the best of our knowledge, the only related modeling study that has accounted for this structure–material interaction is the 1D model of Bartel et al. [32]. In their model, however, instabilities do not originate from a softening-type intrinsic material response but are rather treated as weak displacement discontinuities that separate the transformed and untransformed material points (indeed, experiments, e.g., [35, 36], have confirmed that the true intrinsic response of NiTi is characterized by a significant softening branch). It should be remarked that recently Xiao and Jiang [37, 38] have acknowledged this structure–material interaction in their simulations, however, their applications did not specifically pertain to the subloop deformation and return-point memory.

In light of the above premise, this work aims to provide a detailed analysis of the phenomenon of return-point memory by accounting for the interaction between propagating instabilities and the degradation of the functional properties of the material. To achieve this, a gradient-enhanced model of pseudoelasticity with functional degradation is developed in this work. The model is formulated within the small-strain theory. The basic structure of the model follows the non-gradient model of pseudoelasticity developed by Stupkiewicz and Petryk [39] and is based on the energy minimization principle. The gradient-enhancement, micromorphic regularization, and thermomechanical coupling are adopted from our previously-developed gradient-enhanced model [40, 41]. This previous model was proven to be capable of reproducing the complex patterns of phase transformation in pseudoelastic NiTi specimens under uniaxial tension [41], including the effect of loading rate and latent heat of transformation on martensite domain formation, and in pseudoelastic NiTi tubes under combined tension–torsion [42, 43]. The main advancement of the model in the present work compared to the previous version lies in the incorporation of permanent inelastic strain and the enrichment of the constitutive equations with functional degradation effects. Consequently, given its ability to treat localization effects via gradient-enhancement and micromrophic regularization, the model can be considered a suitable tool for addressing problems where both cyclic loading and transformation localization are at play.

In what follows, we first introduce the model in Section 2. The model is employed to analyze the problem of subloop deformation in NiTi wire and strip under uniaxial tension. The corresponding results are presented and discussed in Section 3. In addition, a simplified version of the model is provided in Appendix A.

2. A small-strain model of pseudoelasticity with functional degradation

The present model falls in the category of phenomenological models. Accordingly, the constitutive relations are tailored, in a simple phenomenological manner, to mimic the pseudoelasticity degradation effects. Since the focus of this study is on the analysis of the return-point memory, which is relevant at the macroscopic scale, a phenomenological description seems to adequately fulfill the intended purpose. In Section 2.1, we introduce the constitutive model in an isothermal format. Subsequently, in Section 2.2, micromorphic regularization, thermomechanical coupling, and finite-element implementation are briefly discussed.

2.1. Constitutive model

We begin the model description by noting that functional fatigue in SMAs is typically attributed to a number of mechanisms. Among them, generation of dislocation slip [4, 44], formation of stabilized martensite [21, 45] and non-transforming austenite [45, 46] are the most likely involved mechanisms. In the present model, a detailed subdivision into the possible mechanisms and their mutual interaction is not attempted, instead, they are unitedly represented by phenomenological evolution equations, and are directly linked to the martensitic phase transformation through the accumulated martensite volume fraction η^{acc}. In line with this notion, the inelastic mechanism responsible for functional degradation is herein denoted as transformation-induced plasticity (TRIP).

The material state at each point is characterized by two quantities, namely the total strain $\boldsymbol{\varepsilon} = \frac{1}{2}\left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^{\mathrm{T}}\right)$, with $\boldsymbol{u}$ as the displacement vector, and the martensite volume fraction η. The total strain is additively decomposed into

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^{\mathrm{e}} + \boldsymbol{\varepsilon}^{\mathrm{t}} + \boldsymbol{\varepsilon}^{\mathrm{p}}, \tag{1}$$

where $\boldsymbol{\varepsilon}^{\mathrm{e}}$ denotes the elastic contribution, $\boldsymbol{\varepsilon}^{\mathrm{t}}$ denotes the martensitic transformation contribution and $\boldsymbol{\varepsilon}^{\mathrm{p}}$ is the permanent strain associated with TRIP. At the same time, it is assumed that during the martensitic transformation a fraction of martensite stabilizes and does not transform back to austenite. Hence, the martensite volume fraction η is split into the reversible part η^{rev} and irreversible part η^{ir}, viz.,

$$\eta = \eta^{\mathrm{rev}} + \eta^{\mathrm{ir}}, \tag{2}$$

and the following inequality constraints hold,

$$0 \leq \eta^{\mathrm{ir}} \leq \eta \leq 1 \quad \Longrightarrow \quad 0 \leq \eta^{\mathrm{rev}} \leq 1 - \eta^{\mathrm{ir}}. \tag{3}$$

The material is in the fully austenitic state when $\eta = \eta^{\mathrm{rev}} = 0$ and is in the fully martensitic state when $\eta = 1$. Nevertheless, once the material starts transforming to martensite from a pristine austenitic state, η^{ir} becomes immediately nonzero, as indicated by Eqs. (4)–(6) below, and thereby, a fully austenitic state will not be recoverable.

It has been repeatedly observed in the experiments that the degradation of pseudoelasticity in conventional polycrystalline NiTi are mostly pronounced during the first tens of cycles, gradually diminishing and eventually reaching saturation as the material passes the so-called shakedown stage, e.g., [22, 47, 48]. In view of this general consensus, we adopt the assumption that both the irreversible volume fraction η^{ir} and the permanent strain $\boldsymbol{\varepsilon}^{\mathrm{p}}$ follow exponential-type evolution laws. Note that this assumption is not unique to the present model and has been exploited in various SMA models that account for functional degradation, e.g., [37, 49–51]. With this assumption in place, we first introduce the accumulated volume fraction η^{acc} as

$$\dot{\eta}^{\mathrm{acc}} = |\dot{\eta}^{\mathrm{rev}}| \quad \Longrightarrow \quad \eta^{\mathrm{acc}} = \int_0^t |\dot{\eta}^{\mathrm{rev}}| \, \mathrm{d}\tau, \tag{4}$$

where the overdot denotes the rate of change of the variable and t denotes the time. The evolution equation for the irreversible volume fraction η^{ir} is then explicitly postulated as

$$\eta^{\mathrm{ir}} = h_{\mathrm{ir}}^{\mathrm{sat}}(1 - \exp(-C_{\mathrm{p}}\eta^{\mathrm{acc}})), \tag{5}$$

which results from the time-integration of the following rate equation (with $\eta^{\mathrm{acc}}\big|_{t=0} = 0$ and $\eta^{\mathrm{ir}}\big|_{t=0} = 0$, as for the initial conditions),

$$\dot{\eta}^{\mathrm{ir}} = h_{\mathrm{ir}}^{\mathrm{sat}} C_{\mathrm{p}} \exp(-C_{\mathrm{p}}\eta^{\mathrm{acc}})\dot{\eta}^{\mathrm{acc}}. \tag{6}$$

Analogously, the evolution equation for the permanent strain $\boldsymbol{\varepsilon}^{\mathrm{p}}$ is postulated as

$$\dot{\boldsymbol{\varepsilon}}^{\mathrm{p}} = \epsilon_{\mathrm{p}}^{\mathrm{sat}} C_{\mathrm{p}} \exp(-C_{\mathrm{p}}\eta^{\mathrm{acc}})\dot{\eta}^{\mathrm{acc}} \boldsymbol{N}_{\mathrm{p}}. \tag{7}$$

In Eqs. (5)–(7), $h_{\mathrm{ir}}^{\mathrm{sat}}$ and $\epsilon_{\mathrm{p}}^{\mathrm{sat}}$ represent the respective saturation values for irreversible volume fraction and permanent strain, C_{p} is the degradation rate, and $\boldsymbol{N}_{\mathrm{p}}$ is the direction tensor which is defined such that the rate of the permanent strain $\dot{\boldsymbol{\varepsilon}}^{\mathrm{p}}$ is aligned with the martensitic transformation strain $\boldsymbol{\varepsilon}^{\mathrm{t}}$, i.e.,

$$\boldsymbol{N}_{\mathrm{p}} = \frac{\boldsymbol{\varepsilon}^{\mathrm{t}}}{\|\boldsymbol{\varepsilon}^{\mathrm{t}}\|}, \quad \|\boldsymbol{\varepsilon}^{\mathrm{t}}\| = \sqrt{\mathrm{tr}(\boldsymbol{\varepsilon}^{\mathrm{t}})^2}. \tag{8}$$

Note that, in view of the definition of the accumulated volume fraction η^{acc}, the variables η^{ir} and $\boldsymbol{\varepsilon}^{\mathrm{p}}$ evolve continuously during both the forward and backward transformations.

Martensitic transformation in SMAs usually exhibits negligible volumetric change [1]. The transformation strain $\boldsymbol{\varepsilon}^{\rm t}$ is therefore assumed to be deviatoric (i.e., $\operatorname{tr}\boldsymbol{\varepsilon}^{\rm t} = 0$). Moreover, since the stress-induced transformation renders the martensite variants to be oriented in the direction of the applied stress, martensite is here considered to appear in a fully-oriented state so that the transformation strain $\boldsymbol{\varepsilon}^{\rm t}$ is defined as a function of the reversible volume fraction $\eta^{\rm rev}$ and the transformation strain of fully-oriented martensite $\bar{\boldsymbol{\varepsilon}}^{\rm t}$,

$$\boldsymbol{\varepsilon}^{\rm t} = \eta^{\rm rev}\bar{\boldsymbol{\varepsilon}}^{\rm t}, \quad \bar{\boldsymbol{\varepsilon}}^{\rm t} \in \bar{\mathcal{P}} = \{\, \bar{\boldsymbol{\varepsilon}}^{\rm t} : \, g(\bar{\boldsymbol{\varepsilon}}^{\rm t}) = 0 \,\}. \tag{9}$$

The set $\bar{\mathcal{P}}$ defines the admissible limit transformation strain tensors characterized by the surface $g(\bar{\boldsymbol{\varepsilon}}^{\rm t}) = 0$ which is expressed in the following form [52],

$$g(\bar{\boldsymbol{\varepsilon}}^{\rm t}) = \left[(-I_2)^{3/2} - bI_3 - cI_4^3\right]^{1/3} - a. \tag{10}$$

In Eq. (10), I_2 and I_3 denote the principal invariants of the limit transformation strain tensor $\bar{\boldsymbol{\varepsilon}}^{\rm t}$ while I_4 denotes a mixed invariant, defined as

$$I_2 = -\frac{1}{2}\operatorname{tr}(\bar{\boldsymbol{\varepsilon}}^{\rm t})^2, \quad I_3 = \det \bar{\boldsymbol{\varepsilon}}^{\rm t}, \quad I_4 = \boldsymbol{m}\cdot\bar{\boldsymbol{\varepsilon}}^{\rm t}\,\boldsymbol{m}, \tag{11}$$

where $\boldsymbol{m}$ is the axis of the transverse isotropy. The parameters a, b and c characterize the shape and size of the surface $g(\bar{\boldsymbol{\varepsilon}}^{\rm t}) = 0$ and are specified as

$$a = \epsilon_{\rm T}\left[\frac{3\sqrt{3}}{4(1+\alpha^3)}\right]^{1/3}, \qquad b = \frac{\sqrt{3}}{6}\frac{9\alpha^3\beta^3 - 7\alpha^3 + 7\beta^3 - 9}{(1+\alpha^3)(1+\beta^3)}, \qquad c = \frac{2\sqrt{3}}{3}\frac{\alpha^3 - \beta^3}{(1+\alpha^3)(1+\beta^3)}, \tag{12}$$

with $\epsilon_{\rm T}$ as the maximum transformation strain in tension, α as the tension–compression asymmetry ratio in the direction along the axis of transverse isotropy (i.e., parallel to $\boldsymbol{m}$), and β as the tension–compression asymmetry ratio in the direction perpendicular to the axis of transverse isotropy (i.e., perpendicular to $\boldsymbol{m}$).

It is noteworthy that the deviatoric nature of the transformation strain $\boldsymbol{\varepsilon}^{\rm t}$ dictates, in accordance with the definition of the direction tensor $\boldsymbol{N}_{\rm p}$, see Eq. (8), that the permanent strain $\boldsymbol{\varepsilon}^{\rm p}$ is also deviatoric. Models within the present context often postulate that the permanent inelastic strain evolves in the direction of stress deviator, e.g., [37, 49, 50]. In the present formulation, it can be easily shown that the stress deviator is perpendicular to the surface $g(\bar{\boldsymbol{\varepsilon}}^{\rm t}) = 0$, see [39], and thereby, the transformation strain $\boldsymbol{\varepsilon}^{\rm t}$ depends on the direction of stress deviator. This, however, does not imply that the transformation strain $\boldsymbol{\varepsilon}^{\rm t}$, and accordingly the permanent strain rate $\dot{\boldsymbol{\varepsilon}}^{\rm p}$, are colinear with the stress deviator.

Another important aspect to highlight is that the accumulation of the irreversible volume fraction $\eta^{\rm ir}$ and its impact on the reversible volume fraction $\eta^{\rm rev}$ cause the magnitude of the transformation strain $\boldsymbol{\varepsilon}^{\rm t}$, which serves as the actual transformation strain measure in the present model, to decrease. However, the surface $g(\bar{\boldsymbol{\varepsilon}}^{\rm t}) = 0$ and so the limit transformation strain $\bar{\boldsymbol{\varepsilon}}^{\rm t}$ remain intact throughout the cyclic transformation. This represents an underlying modeling assumption in the present framework regarding the interaction between phase transformation and cyclic degradation. It reflects the notion that the inherent characteristics of the transformation strain are not affected during the cyclic degradation. Instead, it is the accumulation of TRIP and the decrease in the amount of transformable (reversible) martensite that lead to the contraction of the maximum attainable transformation strain.

We now elaborate on the Helmholtz free energy function and the dissipation potential, both customized to incorporate the degradation effects. Assuming an isothermal process, the Helmholtz free energy ϕ is composed of the following contributions: the chemical energy $\phi_{\rm chem}$, the elastic strain energy $\phi_{\rm el}$, the austenite–martensite interaction energy $\phi_{\rm int}$, the energy of the diffuse interface $\phi_{\rm grad}$, and the energy contribution $\phi_{\rm deg}$ related to the pseudoelasticity degradation, i.e.,

$$\phi(\boldsymbol{\varepsilon},\bar{\boldsymbol{\varepsilon}}^{\rm t},\boldsymbol{\varepsilon}^{\rm p},\eta^{\rm rev},\nabla\eta^{\rm rev},\eta^{\rm ir}) = \phi_{\rm chem}(\eta^{\rm rev},\eta^{\rm ir}) + \phi_{\rm el}(\boldsymbol{\varepsilon},\bar{\boldsymbol{\varepsilon}}^{\rm t},\boldsymbol{\varepsilon}^{\rm p},\eta^{\rm rev}) + \phi_{\rm int}(\eta^{\rm rev}) + \phi_{\rm grad}(\nabla\eta^{\rm rev}) + \phi_{\rm deg}(\eta^{\rm rev},\eta^{\rm ir}). \quad (13)$$

Among the contributions to the Helmholtz free energy ϕ, only $\phi_{\rm deg}$ is specific to the present model. The remaining contributions are rather standard and adhere to the non-cyclic model of pseudoelasticity [39–41] and are formulated as

$$\phi_{\rm chem}(\eta^{\rm rev},\eta^{\rm ir}) = (1-\eta)\phi_0^{\rm a} + \eta\phi_0^{\rm m} = \phi_0^{\rm a} + \Delta\phi_0\eta, \quad (14)$$

$$\phi_{\rm el}(\boldsymbol{\varepsilon},\bar{\boldsymbol{\varepsilon}}^{\rm t},\boldsymbol{\varepsilon}^{\rm p},\eta^{\rm rev}) = \mu\,{\rm tr}(\boldsymbol{\varepsilon}^{\rm e}_{\rm dev})^2 + \frac{1}{2}\kappa({\rm tr}\,\boldsymbol{\varepsilon}^{\rm e})^2, \quad \boldsymbol{\varepsilon}^{\rm e} = \boldsymbol{\varepsilon} - \eta^{\rm rev}\bar{\boldsymbol{\varepsilon}}^{\rm t} - \boldsymbol{\varepsilon}^{\rm p}, \quad (15)$$

$$\phi_{\rm int}(\eta^{\rm rev}) = \frac{1}{2}H_{\rm int}(\eta^{\rm rev})^2, \quad (16)$$

$$\phi_{\rm grad}(\nabla\eta^{\rm rev}) = \frac{1}{2}G\nabla\eta^{\rm rev}\cdot\nabla\eta^{\rm rev}. \quad (17)$$

Here, $\Delta\phi_0 = \phi_0^{\rm m} - \phi_0^{\rm a}$ is the phase transformation chemical energy, μ is the elastic shear modulus and is calculated via applying the Reuss averaging scheme based on the total volume fraction η to the shear moduli of austenite $\mu_{\rm a}$ and martensite $\mu_{\rm m}$ (i.e., $1/\mu = (1-\eta)/\mu_{\rm a} + \eta/\mu_{\rm m}$), κ is the elastic bulk modulus (assumed constant), $H_{\rm int}$ is the parameter that characterizes the material response (softening- or hardening-type) within the transformation regime, and $G > 0$ is the gradient energy coefficient. Note that the parameter $H_{\rm int}$ can be adapted such that it reflects a loading-dependent material response (typically, a softening-type response in tension and hardening-type response in compression), e.g., [42]. However, for simplicity, $H_{\rm int}$ is here considered as a constant parameter. Given that the simulations in this study involve predominantly tensile loading, see Section 3, this simplification does not pose a serious limitation. Note also that the interaction energy $\phi_{\rm int}$ is a quadratic function of the volume fraction $\eta^{\rm rev}$, resulting in a tri-linear intrinsic stress–strain response, as illustrated in Fig. 2. This choice is also made for simplicity and can be readily adapted to more complex functions to achieve a more realistic response [43].

On the other hand, the degradation contribution $\phi_{\rm deg}$ takes the following form

$$\phi_{\rm deg}(\eta^{\rm rev},\eta^{\rm ir}) = A_{\rm deg}\eta^{\rm ir}\eta^{\rm rev} + \frac{1}{2}H_{\rm deg}\eta^{\rm ir}(\eta^{\rm rev})^2, \quad (18)$$

where $A_{\rm deg}$ and $H_{\rm deg}$ represent the degradation parameters. The contribution $\phi_{\rm deg}$ is specifically tailored to address two primary effects of pseudoelasticity degradation: it accounts for the reduction of the transformation-onset stress (described by the term $A_{\rm deg}\eta^{\rm ir}\eta^{\rm rev}$) and the conversion of the mechanical response towards a hardening-type response (described by the term $\frac{1}{2}H_{\rm deg}\eta^{\rm ir}(\eta^{\rm rev})^2$). In line with the evolution of $\eta^{\rm ir}$, Eq. (5), both effects progress exponentially. Note that the approach of incorporating the cyclic degradation effects into the free energy function has been also used in other SMA models in the literature, e.g., [53, 54].

Finally, a rate-independent dissipation potential is adopted in the following form

$$D(\dot{\eta}^{\rm rev},\eta^{\rm acc}) = f_{\rm c}(\eta^{\rm acc})|\dot{\eta}^{\rm rev}|, \tag{19}$$

where $f_{\rm c}(\eta^{\rm acc})$, which is called the critical thermodynamic driving force, controls the width of the hysteresis loop in the stress–strain response. To capture the decrease in the hysteresis loop area (i.e., the dissipated energy) during the cyclic transformation, the parameter $f_{\rm c}$ is defined in relation to the accumulated volume fraction $\eta^{\rm acc}$. Similar to the permanent strain $\boldsymbol{\varepsilon}^{\rm p}$ and the irreversible volume fraction $\eta^{\rm ir}$, Eqs. (5)–(7), $f_{\rm c}$ evolves exponentially as follows

$$f_{\rm c}(\eta^{\rm acc}) = f_{\rm c}^{\rm fin} + (f_{\rm c}^{\rm ini} - f_{\rm c}^{\rm fin})\exp(-C_{\rm f}\,\eta^{\rm acc}), \tag{20}$$

where $f_{\rm c}^{\rm ini}$ and $f_{\rm c}^{\rm fin}$ represent, respectively, the initial and final values of $f_{\rm c}$, and $C_{\rm f}$ denotes the corresponding evolution rate.

To formulate the incremental energy minimization problem, we derive the time-discrete version of the constitutive equations by employing the backward Euler scheme. Having known the variables related to the previous time step t_n, the variables related to the current time step $t_{n+1} = t_n + \Delta t$ are sought. We begin by approximating the incremental evolution equation for the irreversible volume fraction $\eta^{\rm ir}$ and the permanent strain $\boldsymbol{\varepsilon}^{\rm p}$,

$$\Delta t\,\dot{\eta}^{\rm ir} \approx \Delta\eta^{\rm ir} = h_{\rm ir}^{\rm sat}C_{\rm p}\exp(-C_{\rm p}\eta^{\rm acc})\Delta\eta^{\rm acc}, \quad \Delta t\,\dot{\boldsymbol{\varepsilon}}^{\rm p} \approx \Delta\boldsymbol{\varepsilon}^{\rm p} = \epsilon_{\rm p}^{\rm sat}C_{\rm p}\exp(-C_{\rm p}\eta^{\rm acc})\Delta\eta^{\rm acc}\boldsymbol{N}_{\rm p}, \tag{21}$$

where

$$\eta^{\rm acc} = \int_0^{t_{n+1}} \Delta\eta^{\rm acc}\,{\rm d}\tau, \quad \Delta\eta^{\rm acc} = |\Delta\eta^{\rm rev}|, \quad \Delta\eta^{\rm rev} = \eta^{\rm rev} - \eta_n^{\rm rev}, \tag{22}$$

with $\eta_n^{\rm rev}$ as the value of the reversible volume fraction from the previous time step t_n. At the same time, the incremental form of the rate-independent dissipation potential is obtained as

$$\Delta D(\Delta\eta^{\rm rev},\eta^{\rm acc}) = f_{\rm c}(\eta^{\rm acc})|\Delta\eta^{\rm rev}|. \tag{23}$$

The solution of the problem is determined via the incremental energy minimization principle [39, 41, 55]. A global incremental potential Π is defined by summing up the increment of the total Helmholtz free energy $\Delta\Phi$ (where $\Phi = \int_B \phi {\rm d}V$), the global dissipation potential $\Delta\mathcal{D}$ (where $\Delta\mathcal{D} = \int_B \Delta D{\rm d}V$) and the potential of the external loads $\Delta\Omega$, and is subsequently minimized with respect to the unknowns $\boldsymbol{u}$, $\bar{\boldsymbol{\varepsilon}}^{\rm t}$ and $\eta^{\rm rev}$, i.e.,

$$\Pi = \Delta\Phi + \Delta\Omega + \Delta\mathcal{D} \;\to\; \min_{\boldsymbol{u},\bar{\boldsymbol{\varepsilon}}^{\rm t},\eta^{\rm rev}} \tag{24}$$

which is subject to the inequality constraints on the reversible volume fraction $\eta^{\rm rev}$, Eq. (3), and

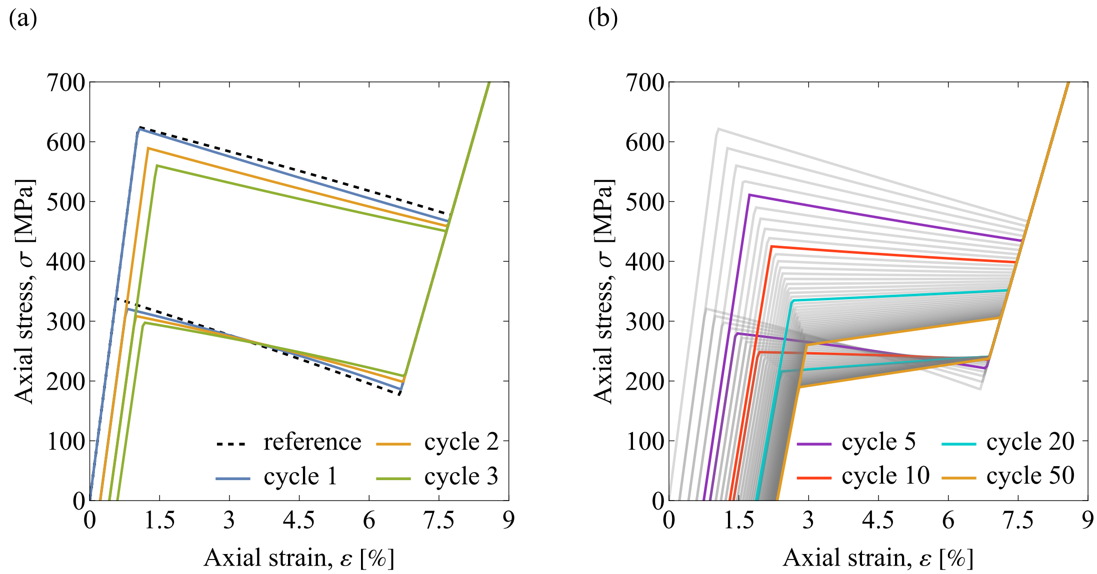


Figure 2: The intrinsic stress–strain response of the model under full-transformation cycles of uniaxial-tension: (a) the first three cycles, and (b) the first 50 cycles. The dashed curve in panel (a), denoted as 'reference', represents the pseudoelastic intrinsic response with no degradation effects. The model parameters adopted to produce the intrinsic responses are the same as those in the main simulations, see Section 3.

to the constraint related to the limit transformation strain surface, Eq. (9). At the same time, η^{ir} and $\boldsymbol{\varepsilon}^{\mathrm{p}}$, which contribute directly to the minimization problem, are explicitly evaluated from Eq. (21). To provide a clearer idea of the structure of the minimization problem and the underlying constitutive behavior of the model, a simplified 1D version of the model is elaborated in Appendix A.

Fig. 2 showcases the intrinsic stress–strain response predicted by the model under cyclic tensile loading. Two cases are highlighted: the pseudoelasticity degradation effects observed within the first three cycles, relevant to the problem of subloop deformation investigated in this study, and the degradation effects observed within 50 cycles, which provides a more holistic view of the model behavior. Note that the material parameters adopted to generate the intrinsic response in Fig. 2 are the same as those adopted in the main simulations in Section 3.

2.2. Further extensions and finite-element implementation

The model presented in Section 2.1 is now enriched with micromorphic regularization and is made thermomechanically coupled. Both extensions have been thoroughly discussed in our previous works [40, 41]. Hence, we only briefly discuss them here.

The purpose of adopting the micromorphic regularization is to facilitate the finite-element implementation of the gradient-enhanced model by restructuring the minimization problem in a way that the constitutive complexities are transferred to the local level (for instance, at the Gauss points) where they can be treated in a more efficient way. To do so, a new degree of freedom $\breve{\eta}$ is introduced and is coupled with the volume fraction η^{rev} through the following penalization term ϕ_{pen} which is added into the Helmholtz free energy function, see Eq. (13),

$$\phi_{\mathrm{pen}}(\eta^{\mathrm{rev}}, \breve{\eta}) = \frac{1}{2}\chi(\eta^{\mathrm{rev}} - \breve{\eta})^2, \tag{25}$$

with χ as the penalty parameter. The gradient energy $\phi_{\rm grad}$, see Eq. (17), is then redefined in terms of the gradient of the new variable $\breve{\eta}$, i.e.,

$$\phi_{\rm grad}(\nabla\breve{\eta}) = \frac{1}{2}G\nabla\breve{\eta}\cdot\nabla\breve{\eta}. \tag{26}$$

Following this modification, the volume fraction $\eta^{\rm rev}$ can be considered as a local quantity and the respective evolution equation can be solved (together with that of $\bar{\boldsymbol{\varepsilon}}^{\rm t}$) at the local level. For further details regarding the micromorphic regularization, interested readers are referred to [56, 57].

To arrive at a thermomechanically-coupled model, two most important couplings are taken into consideration. First, the chemical energy $\phi_{\rm chem}$, Eq. (14), is extended to reflect the effect of temperature on the mechanical response (the Clausius–Clapeyron relation), i.e.,

$$\phi_0(\eta^{\rm rev},\eta^{\rm ir},T) = \phi_0^{\rm a}(T) + \Delta\phi_0(T)\eta, \quad \Delta\phi_0(T) = \Delta s^*(T-T_{\rm t}), \tag{27}$$

where Δs^* represents the transformation entropy change, T is the temperature, and $T_{\rm t}$ is the transformation equilibrium temperature. Next, the internal heat source $\dot{R}$ is defined to encompass the latent heat of transformation and the heat release by mechanical dissipation, viz.,

$$\dot{R} = \Delta s^* T\dot{\eta}^{\rm rev} + f_{\rm c}(\eta^{\rm acc})|\dot{\eta}^{\rm rev}|. \tag{28}$$

Eq. (28) is then introduced into the (isotropic) heat conduction equation

$$\varrho_0 c\dot{T} + \nabla\cdot\boldsymbol{Q} = \dot{R}, \quad \boldsymbol{Q} = -K\nabla T, \tag{29}$$

where $\boldsymbol{Q}$ is the heat flux, $\varrho_0 c$ is the specific heat, and the scalar K is the heat conduction coefficient. It follows from Eq. (28) that the internal heat generation is influenced during the cyclic phase transformation. This influence is manifested by both the latent heat of transformation and the mechanical dissipation and operates through the reversible volume fraction $\eta^{\rm rev}$ and the hysteresis parameter $f_{\rm c}$, cf. Eqs. (3) and (20).

The full thermomechanically-coupled model comprises three global unknown fields: the displacement $\boldsymbol{u}$, the micromorphic variable $\breve{\eta}$ and the temperature T; and two local unknown variables: the reversible volume fraction $\eta^{\rm rev}$ and the limit transformation strain $\bar{\boldsymbol{\varepsilon}}^{\rm t}$. Recall that the irreversible volume fraction $\eta^{\rm ir}$ and the permanent strain $\boldsymbol{\varepsilon}^{\rm p}$ are explicitly integrated by using Eq. (21). The finite-element discretization of the displacement field $\boldsymbol{u}$ is performed by using 20-noded quadratic hexahedral (Serendipity) elements with reduced Gauss integration rule ($2\times2\times2$). On the other hand, 8-noded linear hexahedral elements with standard Gauss integration rule ($2\times2\times2$) are used for $\breve{\eta}$ and T. For the 2D axisymmetric wire problem discussed in Section 3.2, the respective discretizations have been done by 8-noded quadratic elements and 4-noded linear elements. The resulting global–local problem is structured as a nested iterative-subiterative scheme and is solved at both the global and local levels by using the Newton method. Notably, a fully-coupled monolithic scheme is adopted so that the problem is solved simultaneously with respect to all unknowns.

It is worth noting that the local minimization problem of $\eta^{\rm rev}$ is non-smooth, in view of the rate-independent dissipation potential, see Eqs. (19) and (23). To address this issue, the augmented Lagrangian method is utilized, which handles adeptly both the non-smoothness of the rate-independent dissipation potential and the inequality constraints on the reversible volume fraction $\eta^{\rm rev}$, i.e., $0 \leq \eta^{\rm rev} \leq 1 - \eta^{\rm ir}$, see Eq. (3). The local problem has an additional constraint to be satisfied, namely the equality constraint of the limit transformation strain surface, $g(\bar{\varepsilon}^{\rm t}) = 0$, see Eq. (9). The latter is addressed by using a standard Lagrange multiplier method. For brevity, the related technical details are not discussed here, see [39].

The model is transformed into a finite-element code using the automatic differentiation tool AceGen [58, 59], thanks to which the residual vector and the tangent matrix are derived automatically, and thereby, the quadratic convergence of the Newton method is ensured. The simulations are carried out in the finite-element environment AceFEM.

3. Simulations

This section is devoted to the analysis of the simulation results. Section 3.1 begins with a presentation of the simulation setup and calibration of the material parameters, and concludes with a brief discussion on the results for NiTi wire under full loading–unloading cycles. Our main modeling study concerns a NiTi specimen subjected to uniaxial tension with subloop loading paths. Two scenarios are explored. First, in line with the experimental study of Tobushi et al. [24], the subloop deformation behavior of a NiTi wire is analyzed, see Section 3.2. The loading program in this scenario encompasses three nested subloops with increasing strain amplitudes, as depicted in Fig. 3(a). As shown later, this setup enables us to reproduce neatly the hierarchical return-point memory. Section 3.2 concludes with a supplementary analysis of the TRIP evolution under a large number of subloops. Next, in Section 3.3, we extend our analysis to a NiTi strip, where we elucidate how the subloop behavior is influenced by the complexity of the pattern of propagating instabilities. This scenario is then examined under two additional loading programs, see Fig. 3(b,c).

3.1. Preliminaries

In all simulations, the loading is exerted in a displacement-control mode at a (constant) low strain rate of 1.67×10^{-4} s^{-1}. The NiTi wire has a diameter of 0.75 mm and a total length of $L_0 = 20$ mm. To facilitate this analysis, the wire is justifiably reduced to a 2D axisymmetric geometry. The corresponding 2D problem is then discretized by a uniform finite-element mesh consisting of equiaxed elements with an edge size of 0.01 mm. This resulted in 76 000 elements and approximately 620 000 degrees of freedom. Meanwhile, the NiTi strip is treated as a full 3D problem. The strip has a thickness of 0.4 mm, a width of 10 mm and a total length of $L_0 = 100$ mm. It is discretized by a uniform mesh consisting of elements with an in-plane edge size of 0.2 mm and a through-thickness size of 0.4 mm (i.e., only one element is used though the thickness). This mesh led to 25 000 elements and nearly 640 000 degrees of freedom. In both problems, the following boundary conditions are imposed. The displacements at the bottom edge of the specimen are fully constrained. At the top edge, the axial displacement δ is prescribed and the lateral displacements are constrained. At the same time, the temperature at both top and bottom edges is set equal to the ambient temperature, i.e., $T = T_0 = 353$ K, which is the actual ambient temperature maintained during the experiment [24]. Finally, the heat convection effect is neglected.

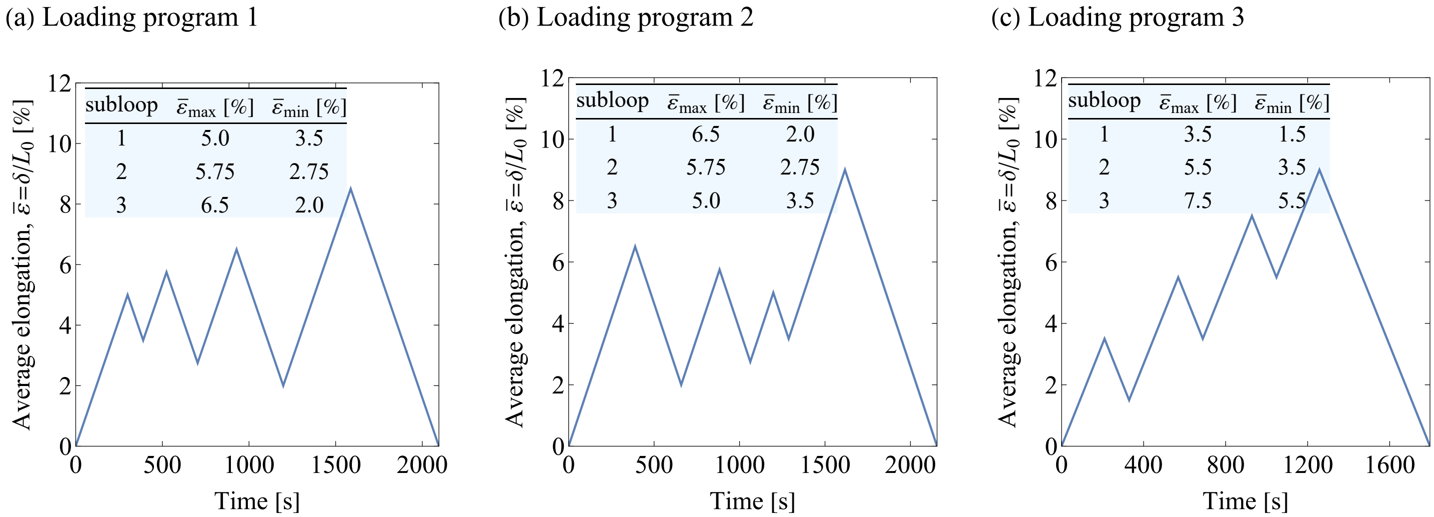


Figure 3: The loading programs used in the simulations. All loading programs represent displacement-controlled uniaxial tension (with a low strain rate of 1.67×10^{-4} s^{-1}) and incorporate three subloops. Loading program 1 consists of three nested subloops with increasing strain amplitudes. Loading program 2 employs subloops in a reverse order compared to loading program 1. Loading program 3 consists of three equally-spaced subloops with a constant strain amplitude.

The model parameters adopted in the simulations are summarized in Tab. 1. Except for the gradient energy parameter G, all the model parameters are identical in the wire and strip problems. The parameter G sets the length-scale associated with the phase transformation front and can be linked to the geometry and micromechanical characteristics [60, 61]. Thus, G takes different values in each problem. To calibrate G, first, an assumption ought to be made regarding the theoretical thickness of the macroscopic interface, λ. Subsequently, G is determined through the analytical relation $G = -H_{\rm int}\lambda^2/\pi^2$, which is derived from the solution of the 1D small-strain model of pseudoelasticity [40]. The identification procedure for the remaining model parameters which are unrelated to TRIP has been thoroughly discussed in our recent study [43], see Section 2.3 and Appendix E therein, and is not repeated here.

Identification of some TRIP-related parameters is guided by the indications obtained from the structural stress–strain response from the experiment, see Fig. 1(a). These include the significant decrease in the level of the upper stress plateau during the hierarchical subloop deformation and the value of the residual strain at the end of the experiment. Accordingly, the parameters $A_{\rm deg} = -45$ MPa, $\epsilon_{\rm r}^{\rm sat} = 0.4\epsilon_{\rm T} = 2.4\%$ (recall that $\epsilon_{\rm T}$ denotes the maximum transformation strain, Eq. (12)) and $C_{\rm p} = 0.05$ have been calibrated to produce similar effects. We, however, acknowledge that there exists a degree of uncertainty in the identification of the remaining parameters, for which we lack definitive experimental evidence. With this in mind, the parameters $H_{\rm deg} = 40$ MPa, $f_{\rm c}^{\rm fin} = 2$ MPa, $C_{\rm f} = C_{\rm p} = 0.05$ and $h_{\rm ir}^{\rm sat} = 0.4$ are selected such that the changes in the stress–strain response under a large number of loading cycles (in particular, as it concerns the transition to a hardening-type response, decrease in the hysteresis loop area and decrease in the extent of the transformation strain) align with the trends observed in the experiments, e.g., [47, 62–64], see also the discussion below. The intrinsic response of the model resulting from the adopted parameters is illustrated in Fig. 2.

Table 1: Model parameters adopted in the simulations.

Category	Parameter		Value
Elasticity	κ	Bulk modulus	130 GPa
	$\mu_{\rm a}$	Shear modulus for austenite	21 GPa
	$\mu_{\rm m}$	Shear modulus for martensite	9 GPa
Martensitic phase transformation	Δs^*	Chemical energy of transformation	0.24 MPa/K
	$T_{\rm t}$	Transformation equilibrium temperature	222 K
	$f_{\rm c}^{\rm ini}$	Hysteresis loop parameter (initial)	10 MPa
	$H_{\rm int}$	Austenite–martensite interaction parameter	-10.5 MPa
	$\epsilon_{\rm T}$	Maximum tensile transformation strain	6%
	α	Tension–compression asymmetry ratio	1.4
	β	Transverse isotropy parameter	1.0
Macroscopic transformation front	G	Gradient energy parameter (wire problem)	0.04 MPa mm^2
	G	Gradient energy parameter (strip problem)	0.4 MPa mm^2
	χ	Micromorphic regularization parameter	100 MPa
Heat transfer	$\varrho_0 c$	Specific heat	2.86 MJ/(m^{3}K)
	K	Heat conductivity	18 W/(m K)
TRIP	$A_{\rm deg}$	Pseudoelasticity degradation parameter	-45 MPa
	$H_{\rm deg}$	Pseudoelasticity degradation parameter	40 MPa
	$\epsilon_{\rm p}^{\rm sat}$	Permanent strain saturation value	0.4 $\epsilon_{\rm T} = 2.4\%$
	$h_{\rm ir}^{\rm sat}$	Irreversible volume fraction saturation value	0.4
	$f_{\rm c}^{\rm fin}$	Hysteresis loop parameter (final)	2 MPa
	$C_{\rm p}$	Degradation rate	0.05
	$C_{\rm f}$	Hysteresis loop degradation rate	0.05

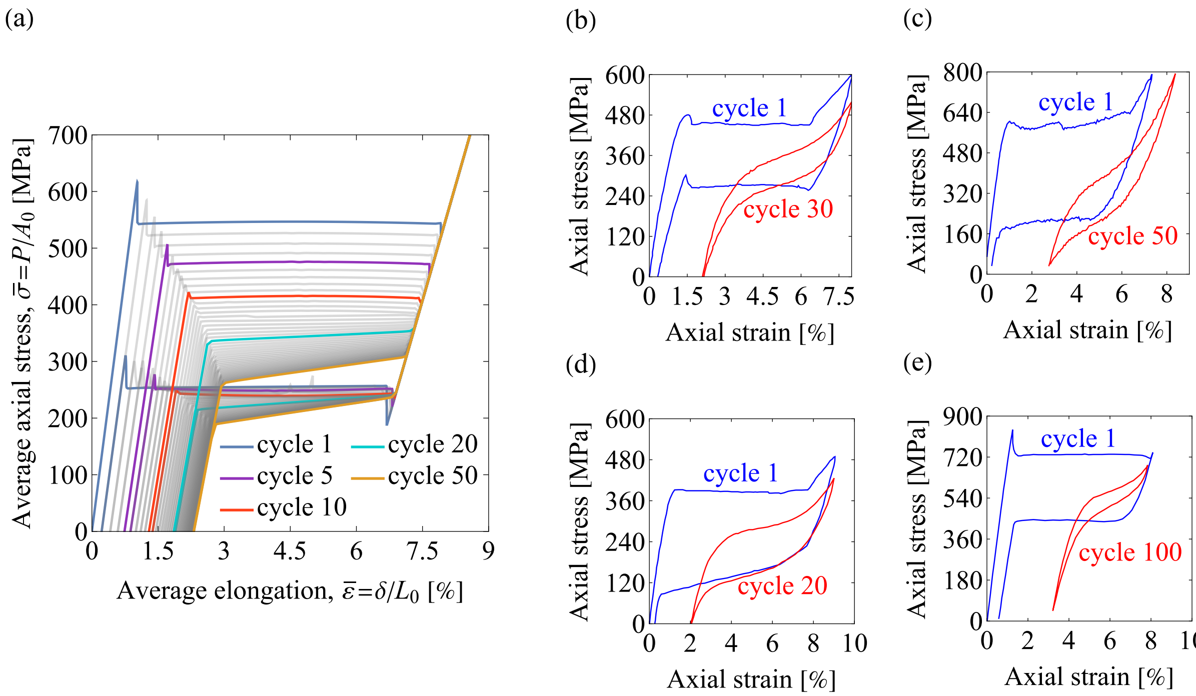


Figure 4: (a) NiTi wire subjected to 50 loading–unloading cycles of uniaxial tension: structural stress–elongation ($\bar{\sigma}$–$\bar{\varepsilon}$) responses. The average axial stress $\bar{\sigma}$ and the average elongation $\bar{\varepsilon}$ are calculated, respectively, as the reaction force P divided by the initial cross-section area A_0, and the axial displacement δ divided by the initial length L_0. The intrinsic responses associated with these structural responses are illustrated in Fig. 2. (b)–(e) Typical cyclic responses of NiTi specimens observed in the experiments, taken from (b) Wang et al. [47], (c) Morin et al. [62], (d) Kan et al. [63], and (e) Šittner et al. [64].

It is worth noting that in all the simulations, as a way to trigger the phase transformation instability, a geometric imperfection in the form of a slight indent is applied to the specimen. The indent is located at a distance equal to the diameter/width of the wire/strip from its lower end.

Before entering into the main analysis of subloop deformation, a simulation is performed for the NiTi wire subjected to 50 loading–unloading cycles of uniaxial tension. Fig. 4 illustrates the structural response of the wire. Here, as well as in the figures in the following subsections, the structural response is represented in terms of the average axial stress $\bar{\sigma} = P/A_0$ versus average elongation (engineering strain) $\bar{\varepsilon} = \delta/L_0$, where P denotes the reaction force and A_0 denotes the initial cross-section area. Recall that δ and L_0 are the axial displacement and the initial length, respectively. As it is evident, the wire undergoes a complete phase transformation within each cycle. Initially, the wire exhibits a localized phase transformation, characterized by a stress drop at the transformation onset and a subsequent stress plateau in the structural stress-elongation ($\bar{\sigma}$–$\bar{\varepsilon}$) response. The localized transformation persists for about 15 cycles. Thereafter, the transformation proceeds in a more homogeneous manner, and the structural response displays a mild hardening. As the number of cycles increases, the slope of the hardening branch also increases. The cyclic behavior captured in the simulation is in a qualitative agreement with the typical cyclic behavior of NiTi specimens observed in experiments [47, 62–64], which underscores the reliability of the simulation results.

3.2. NiTi wire subjected to subloop deformation

The results pertaining to the subloop behavior of the NiTi wire are presented in Figs. 5 and 6. The phase transformation evolution in Fig. 5(a) and TRIP evolution in Fig. 5(b) are displayed via, respectively, the distribution of the reversible volume fraction η^{rev} and irreversible volume fraction η^{ir}. Note that, for a more natural visualization, the results of the 2D axisymmetric wire are post-processed and presented in a 3D configuration. As anticipated, the transformation initiates at the position of the geometric imperfection. Throughout the entire loading stage of the global cycle (hereinafter, to avoid confusion with the subloops, we use the term 'global'), the transformation maintains a single propagating front. Interestingly, while the front appears to be a flat (and visibly diffuse) interface in the 3D-wire configuration, e.g., [65], it takes on a spherical-shaped appearance (or 'cone-shaped' as described in [60, 66]), as can be conceived from the corresponding pattern in the axisymmetric planes (not shown here). During the global unloading, the backward transformation commences from the wire's central part. As shown in Fig. 5(b) and discussed below, the highest amount of irreversible volume fraction η^{ir}, thus the highest TRIP, is accumulated within the central part, making it a favorable site for the nucleation of the austenitic band. At the same time, due to a slight asymmetry in the distribution of η^{ir} with respect to the wire's midpoint, the two evolved fronts do not propagate concurrently. More specifically, first, the top front reaches the boundary and annihilates, which manifests as an abrupt stress rise in the structural stress-elongation ($\bar{\sigma}$–$\bar{\varepsilon}$) response, occurring at an average elongation of about $\bar{\varepsilon} = 4\%$ (see Fig. 6). Subsequently, the bottom front follows suit.

Within each subloop path, the front retreats downward during unloading and advances upward during reloading. This cyclic movement prompts the material points inside the front's sweeping zone to undergo backward-then-forward transformation, and thereby, gives rise to the accumulation of TRIP within the sweeping zone, while the material points beyond it remain unaffected. Note that the loading program adheres to a fixed nominal mean strain (set at $\bar{\varepsilon} = 4.5\%$, which corresponds to the front's proximity to the wire's midpoint) but an increasing strain amplitude, see Fig. 3(a). Thus, the sweeping zone expands successively from subloop 1 to subloop 3, and at the same time, the sweeping zone of each subloop encompasses that of the previous one. This therefore results in the highest concentration of TRIP within the central part of the wire and its step-wise decreasing trend as it moves away from it, as can be clearly seen in Fig. 5(b).

The hierarchical return-point memory, which is an outcome of the cyclic traversal of the front across the boundaries of the swept zones, is correctly reproduced in the structural stress–elongation response in Fig. 6. The reproduced feature is in a reasonable agreement with the experimental result of Tobushi et al. [24], see Fig. 1 and the accompanying discussion. In view of the exponential nature of the pseudoelasticity degradation effects, the reduction in the level of the upper stress plateau is at the highest within the first level of hierarchy (of about $\Delta\bar{\sigma} = 21$ MPa) and diminishes to its lowest within the last level of hierarchy (of about $\Delta\bar{\sigma} = 17$ MPa). It is worth noting that in this scenario, where the strain rate corresponds to nearly isothermal conditions (i.e., the temperature variation lies within the range of -2 K to 2 K), the stress, upon reaching the return-point, appears to catch up closely with the corresponding stress plateau before applying the subloop. As shown in Section 3.3 and also observed in the experiments [25, 26], such a close catching up does not occur when thermal effects are at play.

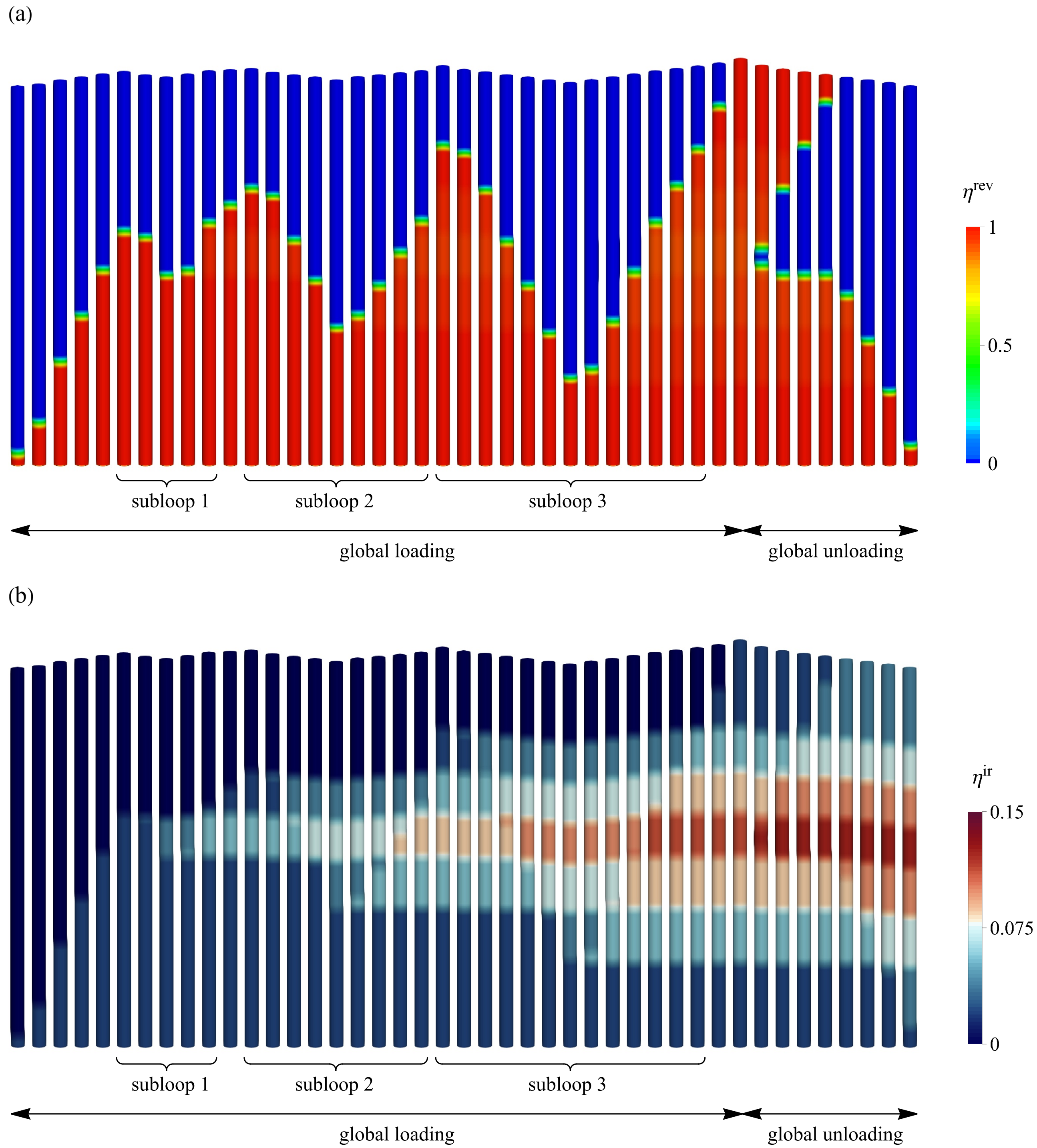


Figure 5: NiTi wire subjected to loading program 1: (a) snapshots of reversible volume fraction η^{rev} illustrating the phase transformation evolution, and (b) snapshots of irreversible volume fraction η^{ir} illustrating the TRIP evolution. For a natural visualization, the axisymmetric wire is presented in full 3D configuration.

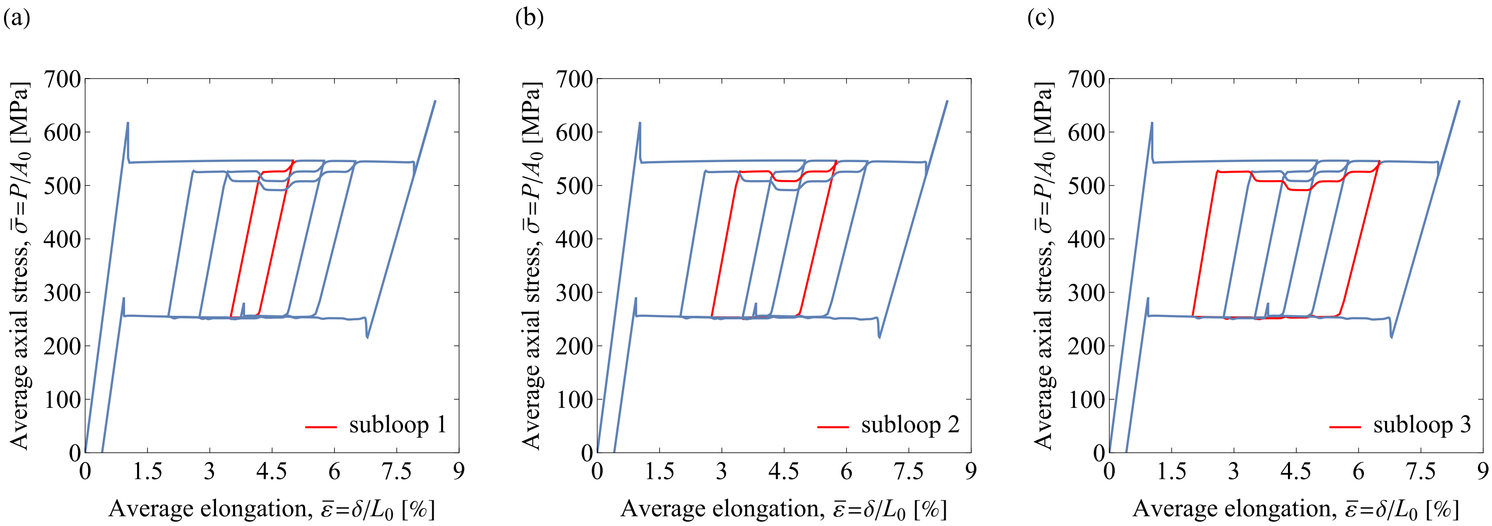


Figure 6: NiTi wire subjected to loading program 1: structural stress–elongation ($\bar{\sigma}$–$\bar{\varepsilon}$) response.

A notable observation from the experimental curve in Fig. 1(a) is the absence of the return-point memory during the unloading stages of the subloops. Instead, the lower stress plateau seems to shift slightly downward from one subloop to the next. Anyway, no attempt was made to adjust the material parameters to replicate the observed behavior, which is, however, present in the results of NiTi strip reported in Section 3.3.

In concluding the discussion in this section, we present the results of a supplementary analysis on the NiTi wire subjected to 12 subloops. The aim of this analysis is to illustrate the evolution of TRIP and subloop deformation behavior over a large number of subloops. The results, as depicted in Fig. 7, follow an expected trend. However, two specific observations deserve further comment. Firstly, the stress plateau in a number of subloops exhibits irregularities, specifically a second stress drop appears ahead of the return-point memory. This is because in these subloops the transformation during subloop reloading does not proceed by the propagation of the existing font. Instead, a second front emerges at the opposite end of the sweeping zone and eventually merges with the original front, thereby, leading to the observed effects. Secondly, as shown in Fig. 7(b), the sweeping zone continuously expands from one subloop to the next. This is explained by the accumulation of TRIP within the sweeping zone, which reduces the amount of transformable martensite. Consequently, since the applied strain amplitude of the subloops is held fixed, the front gradually moves towards the untransformed segments of the wire to compensate for the reduced transformation.

3.3. NiTi strip subjected to subloop deformation

We begin this section by analyzing the NiTi strip under loading program 1. The primary aim is to examine the subloop behavior in a notably more involved scenario than the NiTi wire discussed earlier, arising mainly from a more complex transformation pattern and heightened thermal effects. The simulation results are presented in Figs. 8 and 9. A quick look at Fig. 9 immediately indicates that the return-point memory is only observable in the trajectories that lead to the global stress plateau, while the hierarchical return-point memory is lost. This is undoubtedly an outcome of

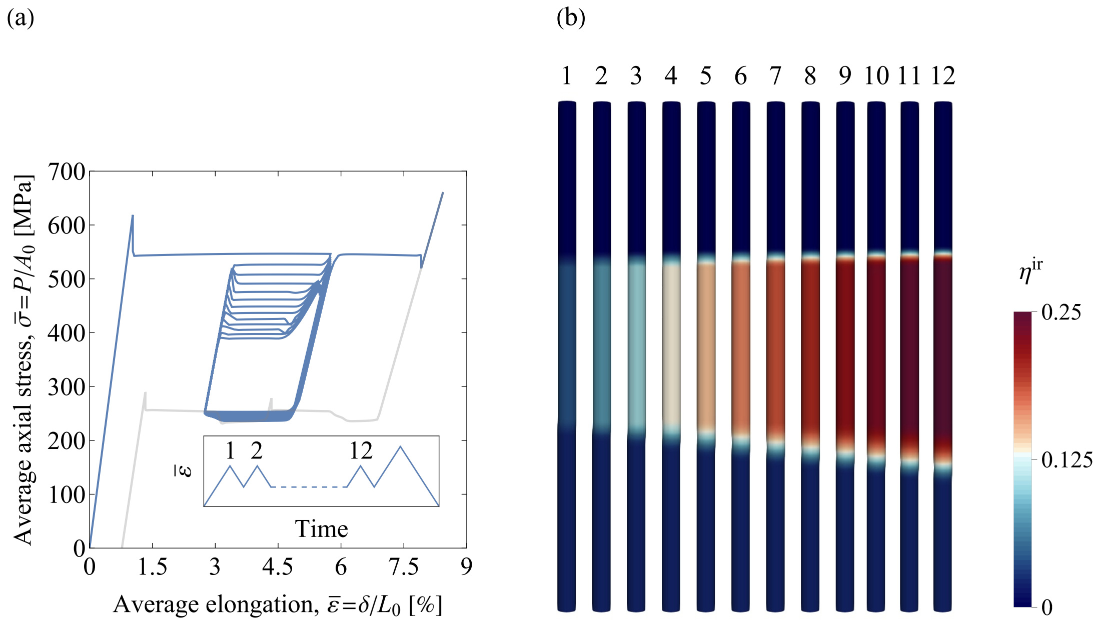


Figure 7: NiTi wire subjected to uniaxial tension with 12 subloop paths: (a) structural stress–elongation ($\bar{\sigma}$–$\bar{\varepsilon}$) response, and (b) snapshots of irreversible volume fraction η^{ir} illustrating the TRIP evolution at the end of each subloop. The inset in panel (a) represents the loading program used for this simulation. The applied subloops have the same mean strain and the same strain amplitude, i.e., with $\bar{\varepsilon}_{\max} = 5.75\%$ and $\bar{\varepsilon}_{\min} = 2.75\%$.

the nontrivial pattern of phase transformation and resulting TRIP distribution within the strip. In contrast to the NiTi wire, where a single phase transformation front remained active during all subloops, the strip features multiple transformation fronts, each presenting a less predictable pattern of activation. Below, we provide a more detailed account of the unfolding events.

The phase transformation initiates with the nucleation of a single martensite band at the location of the geometric imperfection. The band is oriented at approximately 54° with respect to the longitudinal axis, in agreement with the experimental observations [14] and theoretical analysis [67], and changes gradually as loading progresses. At an average elongation of about $\bar{\varepsilon} = 3\%$, another martensite band emerges at the opposite end, and henceforth, the two transformation fronts propagate towards each other. This non-synchronous double nucleation has been commonly observed in the experiment of NiTi specimens at relatively low strain rates, e.g., [14, 68, 69]. Within subloop 1, the two fronts exhibit a short back-and-forth movement, manifesting a clear return-point memory in the structural stress–elongation response in Fig. 9(a). During subloop 2, not only the hitherto active fronts but also the fronts near the boundaries become engaged in the transformation evolution. As a consequence, TRIP is induced via all fronts. This behavior is reflected in the structural response which takes on an irregular appearance characterized by few sudden stress changes, thus spoiling the return-point memory in the inner part (Fig. 9(b)). A similar process recurs within subloop 3, albeit with a more complex phase transformation evolution during the reloading stage and also more distinct stress events in the structural response. During the global unloading, the backward transformation proceeds predominantly in a criss-cross mode, which persists until an average elongation of about $\bar{\varepsilon} = 3\%$. Subsequently, the fronts reconfigure into sharp inclined interfaces that move towards each other until the complete annihilation of the (reversible)

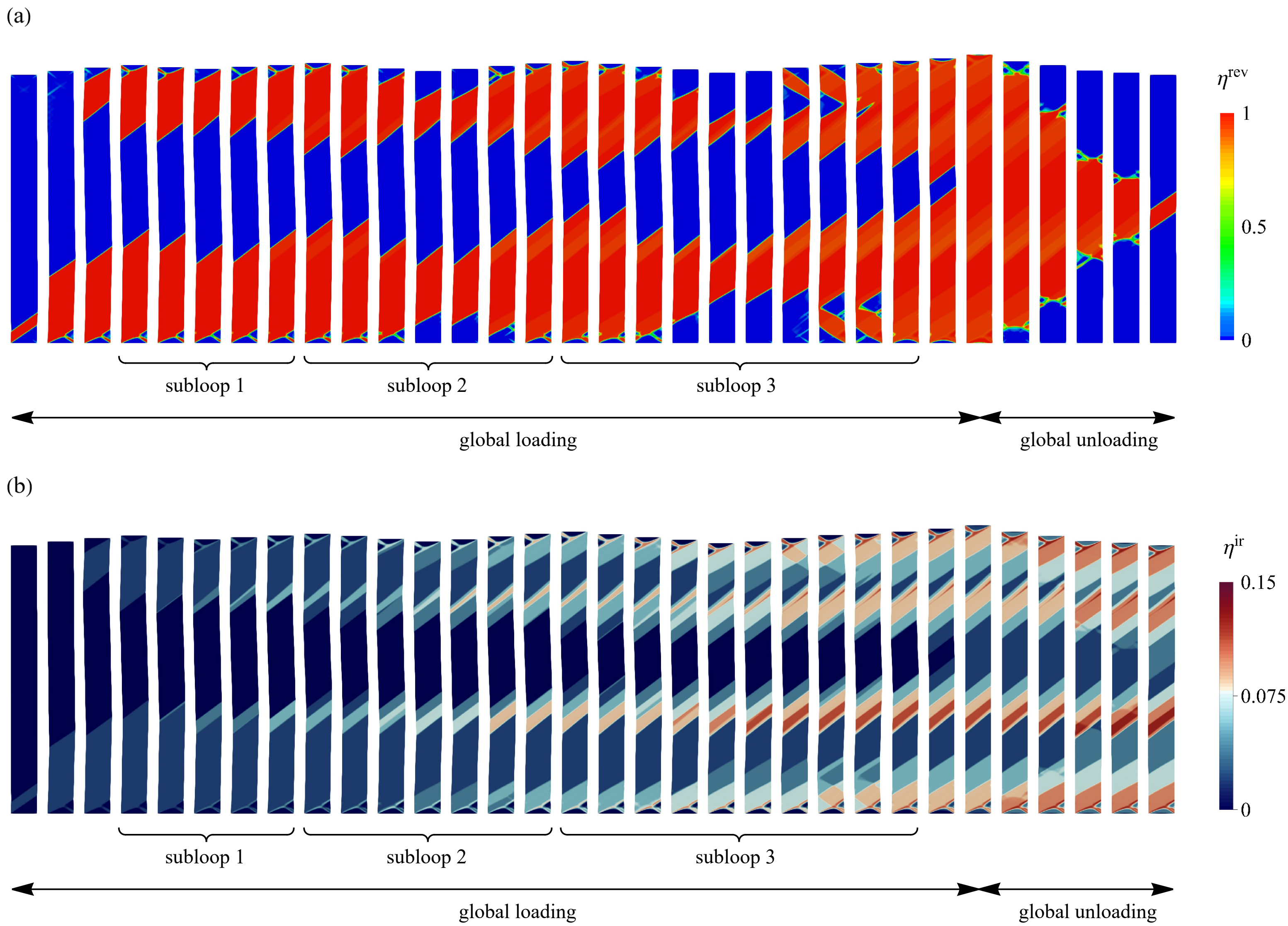


Figure 8: NiTi strip subjected to loading program 1: (a) snapshots of reversible volume fraction η^{rev} illustrating the phase transformation evolution, and (b) snapshots of irreversible volume fraction η^{ir} illustrating the TRIP evolution.

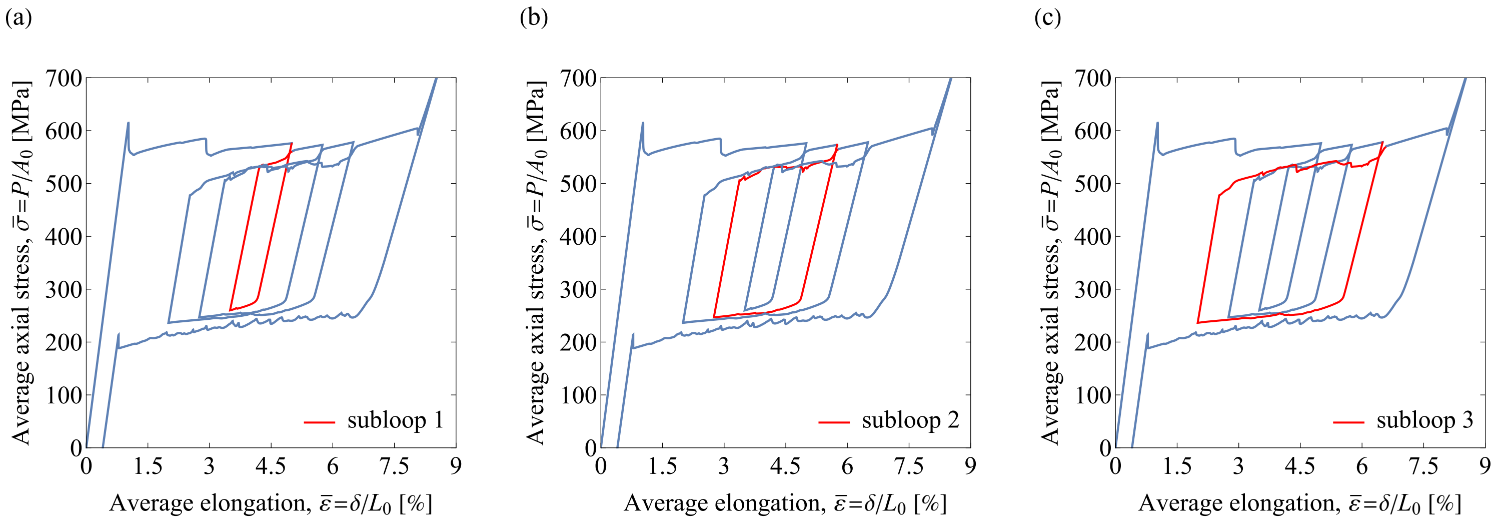


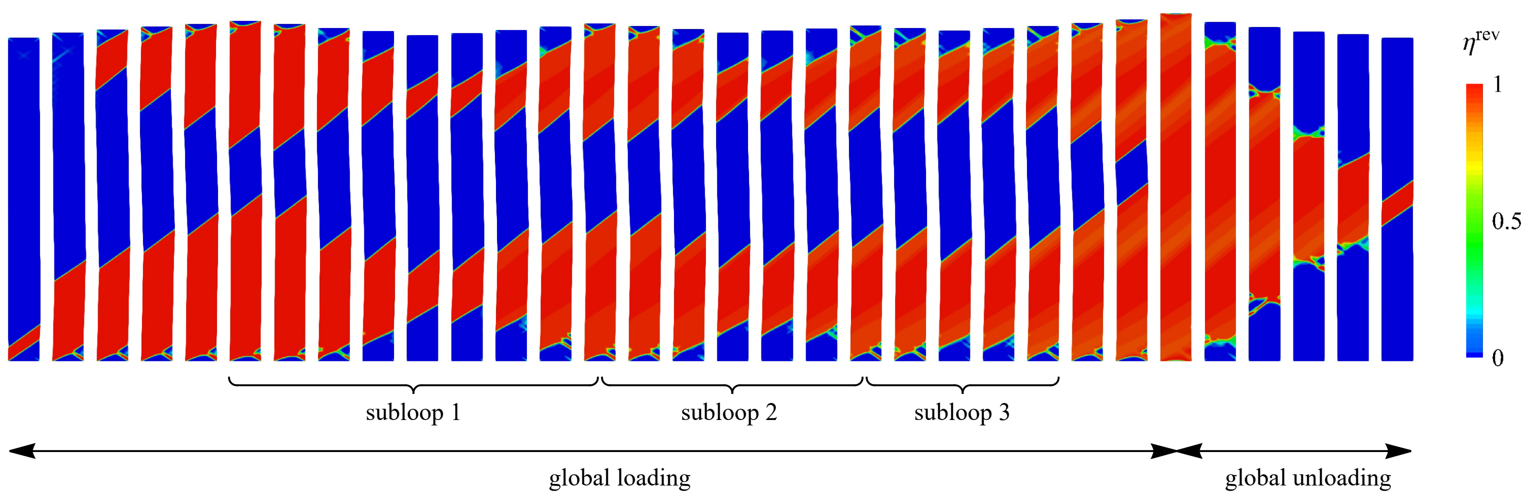
Figure 9: NiTi strip subjected to loading program 1: structural stress–elongation ($\bar{\sigma}$–$\bar{\varepsilon}$) response.

martensite domain. The reconfiguration of the fronts between criss-cross mode and sharp interfaces has been also observed in experimental and previous modeling studies, e.g., [36, 70, 71]. It is worth remarking that the distribution of the irreversible volume fraction η^{ir} within the entire strip at the end of the global unloading remains consistent with that at the end of the global loading, while its magnitude increases uniformly.

Upon inspecting the return-point memory in Fig. 9, a slight difference can be noticed concerning the level of the global stress plateau before and after a subloop path. This difference stems from the thermal effects. Specifically, compared to the NiTi wire, a more pronounced temperature variation is produced across the specimen during the forward transformation (for instance, of about 10 K immediately before subloop 1), resulting in a more visible thermal hardening that sustains a higher stress for the propagation of the front. Within the subloop path, the transformation latent heat is initially absorbed during the backward transformation (self-cooling) and is subsequently released when the forward transformation resumes (self-heating). Accordingly, as the front reaches the pristine material, the temperature variation across the specimen is reduced compared to the state before the subloop (for instance, of about 5 K immediately after subloop 1). Thereby, thermal hardening diminishes, necessitating a lower stress for interface propagation. Note also that as a result of the cyclic transformation of the material points, and thus the accumulation of irreversible martensite, a smaller martensite volume fraction is transformed during the subloop reloading compared to the state before the subloop, and this contributes to the reduction of the latent heat generation [20].

We now proceed with the analysis of the NiTi strip under two additional loading programs, one consisting of nested subloops with decreasing strain amplitudes, i.e., subloops are applied in a reverse order compared to loading program 1, and the other consisting of three equally-spaced distinct subloops with a constant strain amplitude, see Fig. 3(b,c). The corresponding results are shown in Figs. 10, 11 and 12. The comparison of the snapshots of the reversible volume fraction η^{rev} in the two additional cases to those of loading program 1 reveals noticeable morphological differences, which are beyond the differences arising solely from the loading-dependent transformation evolution pathways.

(a)



(b)

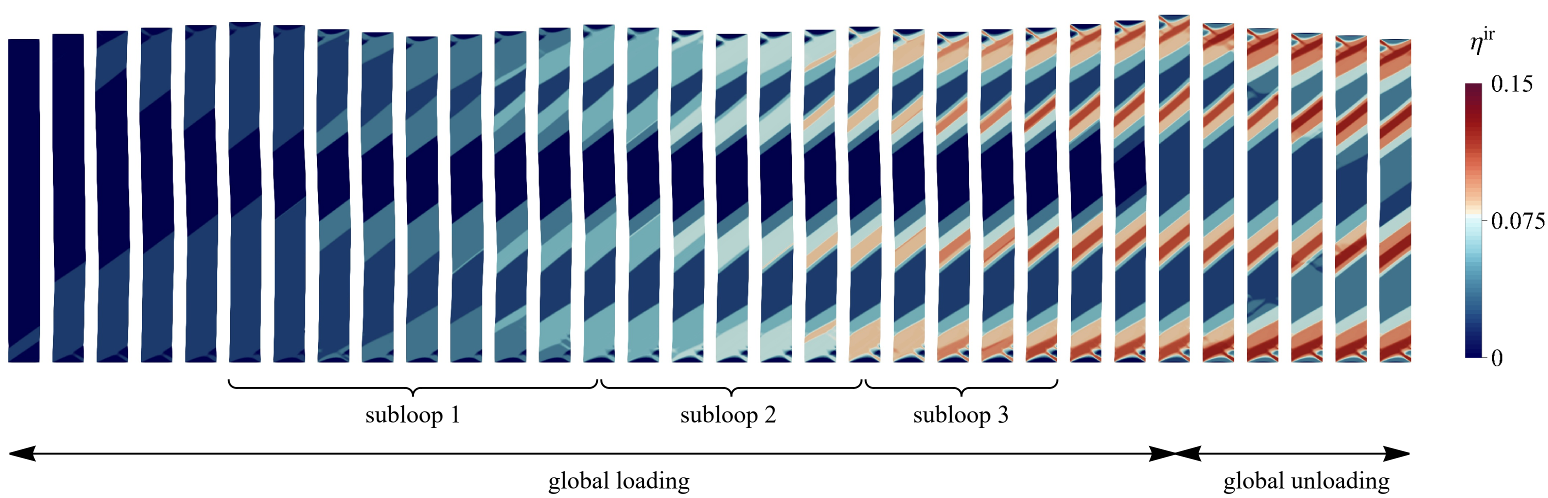


Figure 10: NiTi strip subjected to loading program 2: (a) snapshots of reversible volume fraction η^{rev} illustrating the phase transformation evolution, and (b) snapshots of irreversible volume fraction η^{ir} illustrating the TRIP evolution.

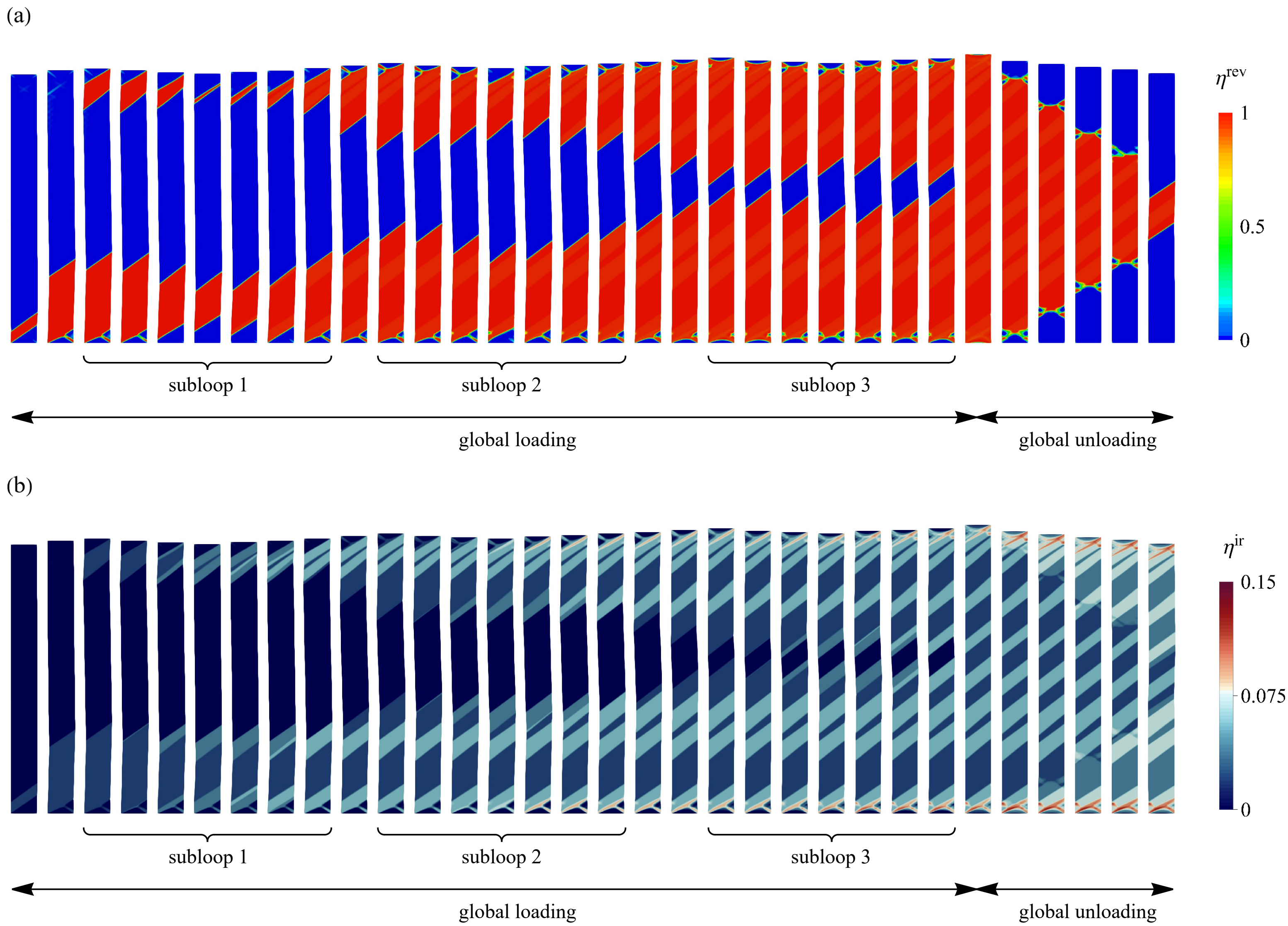


Figure 11: NiTi strip subjected to loading program 3: (a) snapshots of reversible volume fraction η^{rev} illustrating the phase transformation evolution, and (b) snapshots of irreversible volume fraction η^{ir} illustrating the TRIP evolution.

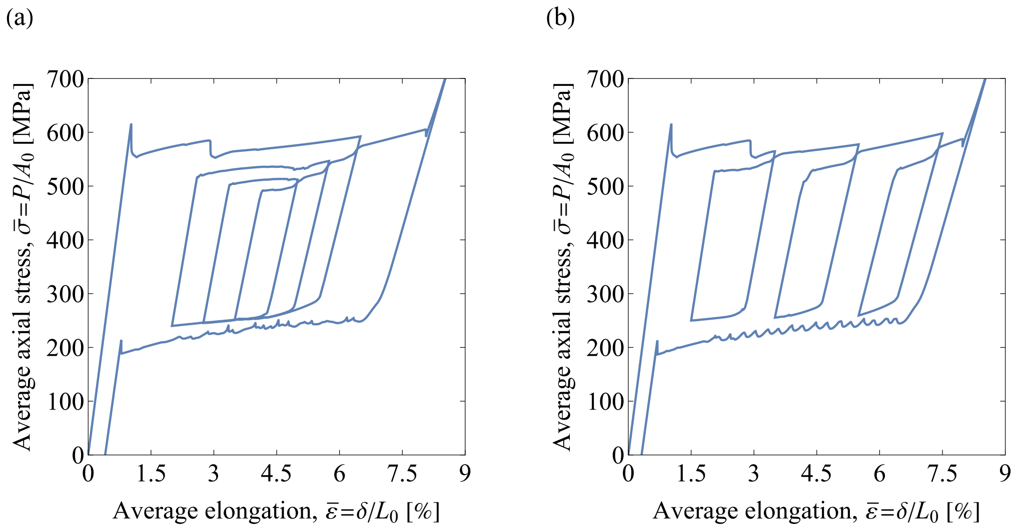


Figure 12: Structural stress–elongation ($\bar{\sigma}$–$\bar{\varepsilon}$) response of NiTi strip subjected to (a) loading program 2 and (b) loading program 3.

The differences mainly concern the varying number of martensite domains formed during the global loading stage and the activation pattern of the fronts within the subloops. More specifically, unlike loading program 1, loading programs 2 and 3 exhibit only two martensite domains during the global loading. In loading program 2, all four fronts remain consistently active within all subloops, resulting in a clear demonstration of the hierarchical return-point memory in the stress–elongation response, as shown in Fig. 12(a). In loading program 3, however, while the involvement of the fronts near the boundaries is eye-catching within subloop 1, overall, the interior fronts are prominently active. In this case, the front sweeping zones in the subloops do not interact with each other (as can be also recognized from the snapshots of η^{ir} in Fig. 11), and the resulting subloops are independent, see Fig. 12(b). During the global unloading, all cases show a similar transformation evolution pattern characterized by two active fronts retracting in a criss-cross manner..

We conclude this discussion by addressing TRIP accumulation within the strip in relation to the loading program. Similar to the martensitic transformation, TRIP exhibits characteristics that are specific to the applied loading program. Given that loading programs 1 and 2 have a reverse arrangement of the subloops but are otherwise identical, one would intuitively expect that the resulting TRIP accumulations, in terms of both the pattern and the intensity, would be the same after applying all the three subloops. A comparison of the snapshots of the irreversible volume fraction η^{ir} (Figs. 8 and 10) indeed confirms that TRIP hotspots in these two cases are located in nearly the same regions, with two hotspots near the boundaries and two within the interior of the strip, corresponding to the regions with the highest activity of the fronts. Yet, minor discrepancies can be observed, particularly concerning the intensity of TRIP within the hotspot regions. On the other hand, loading program 3 demonstrates a rather distinct TRIP accumulation characterized by several regions with mild intensity within the interior and localized hotspots near the boundaries. As previously noted, this particular TRIP distribution results from the lack of interaction among the fronts sweeping zones of the independent subloops. As a summary of this discussion, Fig 13 compares the distribution of η^{ir} along the entire length of the strip for various loading programs.

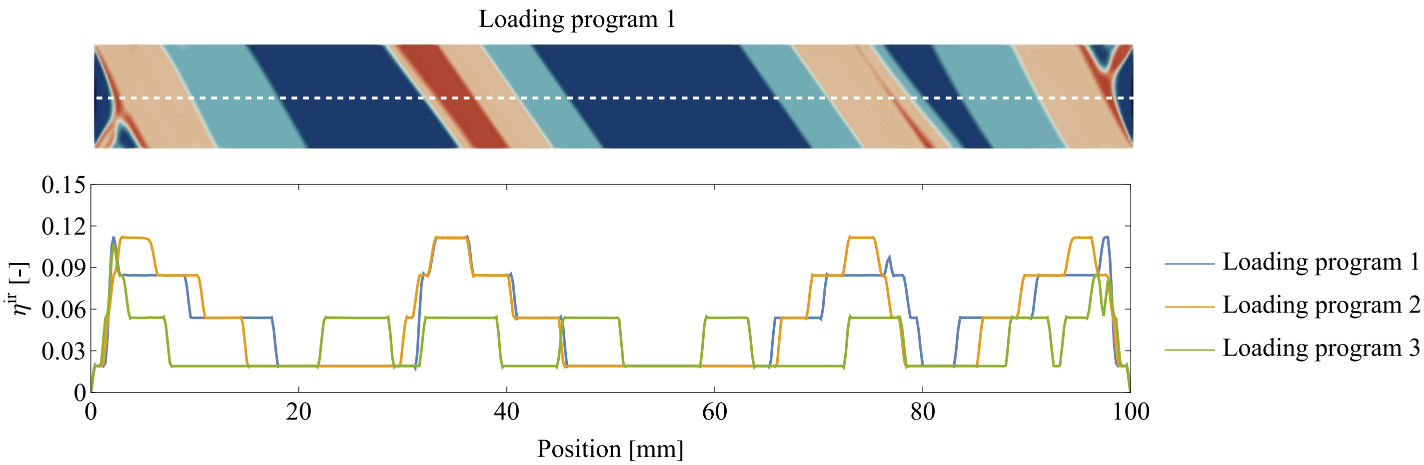


Figure 13: Distribution of the irreversible volume fraction η^{ir} along the entire length of the strip (taken in the reference configuration) at the end of the global loading stage. The graphs correspond to the midsection of the strip, as indicated by the white dashed curve overlaid on the snapshot.

4. Conclusion

The phenomenon of return-point memory that appears during the subloop deformation of pseudoelastic SMA is an outcome of the interaction between the structural instabilities of phase transformation and the degradation of functional properties. It seems that this crucial aspect has been generally overlooked in existing modeling approaches. The goal of our study is to demonstrate this structure–material interaction by modeling the phenomenon of return-point memory. To achieve this, we have developed a gradient-enhanced model of pseudoelasticity. The developed model represents an advancement over previous versions [39–41], extending the constitutive description to incorporate pseudoelasticity degradation. The capabilities of the model in reproducing the essential aspects of pseudoelasticity degradation have been shown for NiTi under cyclic uniaxial tension.

We examine an illustrative example of a NiTi wire subjected to nearly isothermal uniaxial tension with nested subloops. The obtained results clearly correlate with the experimental observations of Tobushi et al. [24], especially regarding the hierarchical return-point memory. The accumulation of TRIP and its distribution during subloop deformation underline the intertwined evolution of inhomogeneous phase transformation and cyclic degradation.

The study is then extended to a more involved scenario of a NiTi strip, where a detailed analysis is performed by examining three different loading programs. The impact of the loading program on the evolution of phase transformation and TRIP has been highlighted through the activation pattern of phase transformation fronts within the subloops, and its implications on the phenomenon of return-point memory have been pointed out. In addition, the results hint at the visible contribution of the thermomechanical coupling effects within the subloops, stemming from the self-cooling/heating process of the transforming material.

Acknowledgement. The authors wish to thank Prof. Stanisław Stupkiewicz for helpful discussions on the model formulation and reviewing the manuscript. This work has been supported by the National Science Centre (NCN) in Poland through the Grant No. 2021/43/D/ST8/02555. Maciej Ryś acknowledges the support from the European Union Horizon 2020 research and innovation program under Grant Agreement No. 857470 and from the European Regional development Fund under the program of the Foundation for Polish Science International Research Agenda PLUS, grant No. MAB PLUS/2018/8, and the initiative of the Ministry of Science and Higher Education 'Support for the activities of Centers of Excellence established in Poland under the Horizon 2020 program' under agreement No. MEiN/2023/DIR/3795. For the purpose of Open Access, the authors have applied a CC-BY public copyright license to any Author Accepted Manuscript (AAM) version arising from this submission.

Appendix A. A simplified 1D demonstration of the model

In this appendix, we derive the governing equation of the transformation stress for a simplified isothermal 1D model. In 1D setting, the model features four variables, namely the total strain $\varepsilon = \nabla u$, the reversible volume fraction η^{rev}, the irreversible volume fraction η^{ir} and the permanent strain ε^{p}. The Helmholtz free energy is thus expressed as follows

$$\phi(\varepsilon, \varepsilon^{\mathrm{p}}, \eta^{\mathrm{rev}}, \eta^{\mathrm{ir}}) = \phi_{\mathrm{chem}}(\eta^{\mathrm{rev}}, \eta^{\mathrm{ir}}) + \phi_{\mathrm{el}}(\varepsilon, \varepsilon^{\mathrm{p}}, \eta^{\mathrm{rev}}) + \phi_{\mathrm{int}}(\eta^{\mathrm{rev}}) + \phi_{\mathrm{deg}}(\eta^{\mathrm{rev}}, \eta^{\mathrm{ir}}) + I(\eta^{\mathrm{rev}}), \quad \text{(A.1)}$$

where the indicator function I pertains to the inequality constraints on the reversible volume fraction η^{rev} ($I = 0$ if $0 < \eta^{\mathrm{rev}} < 1 - \eta^{\mathrm{ir}}$ and $I = \infty$ otherwise). Note that the gradient energy associated with the austenite–martensite diffuse interface, ϕ_{grad}, is disregarded here.

The elastic strain energy ϕ_{el} is formulated as

$$\phi_{\mathrm{el}}(\varepsilon, \varepsilon^{\mathrm{p}}, \eta^{\mathrm{rev}}) = \frac{1}{2} E(\varepsilon - \varepsilon^{\mathrm{t}} - \varepsilon^{\mathrm{p}})^2, \quad \varepsilon^{\mathrm{t}} = \eta^{\mathrm{rev}} \epsilon_{\mathrm{T}}, \quad \text{(A.2)}$$

where E is the Young's modulus (for simplicity, E is assumed constant and independent of η) and the constant ϵ_{T} is the maximum transformation strain. The remaining components of the free energy, as well as the dissipation potential, are identical to those of the general 3D model, see Eqs. (14), (16), (18) and Eq. (23). Moreover, the evolution equations for the permanent strain ε^{p} and the irreversible volume fraction η^{ir} are postulated as (cf. Eqs. (5)–(7))

$$\eta^{\mathrm{ir}} = h_{\mathrm{ir}}^{\mathrm{sat}}(1 - \exp(-C_{\mathrm{p}} \eta^{\mathrm{acc}})), \quad \varepsilon^{\mathrm{p}} = \epsilon_{\mathrm{p}}^{\mathrm{sat}}(1 - \exp(-C_{\mathrm{p}} \eta^{\mathrm{acc}})). \quad \text{(A.3)}$$

For a given total strain ε, the volume fraction η^{rev} can be determined by minimizing the local potential $\pi = \Delta\phi + \Delta D$, see Eq. (24). It is immediate to see that the local potential π is non-smooth, due to the presence of the rate-independent dissipation ΔD and the indicator function I. In line with [40], the minimization of π with respect to η^{rev} is written as a differential inclusion, given by

$$f_{\eta^{\mathrm{rev}}} \in \partial_{\eta^{\mathrm{rev}}} \bar{D}(\eta^{\mathrm{rev}}, \eta^{\mathrm{acc}}) \quad \text{(A.4)}$$

where $\bar{D} = \Delta D + I$ encompasses the non-smooth components of π and $f_{\eta^{\mathrm{rev}}}$ is the thermodynamic driving force associated with η^{rev} and is expressed as

$$f_{\eta^{\mathrm{rev}}} = -\left(\frac{\partial \phi}{\partial \eta^{\mathrm{rev}}} + \frac{\partial \phi}{\partial \varepsilon^{\mathrm{p}}} \frac{\partial \varepsilon^{\mathrm{p}}}{\partial \eta^{\mathrm{rev}}} + \frac{\partial \phi}{\partial \eta^{\mathrm{ir}}} \frac{\partial \eta^{\mathrm{ir}}}{\partial \eta^{\mathrm{rev}}} \right). \quad \text{(A.5)}$$

During the forward/backward transformation, i.e., when the bound constraints are inactive, the inclusion (A.4) yields

$$f_{\eta^{\text{rev}}} = \pm f_{\text{c}}, \tag{A.6}$$

and gives the following equation for the transformation stress $\sigma^{\text{t}}_{\pm}$ (σ^{t}_{+} for the forward transformation and σ^{t}_{-} for the backward transformation),

$$\sigma^{\text{t}}_{\pm} = \frac{\Delta\phi_0 k_1 \pm f_{\text{c}} + H_{\text{int}}\eta^{\text{rev}} + A_{\text{deg}}k_2 + H_{\text{deg}}\eta^{\text{rev}}k_3}{k_4}, \tag{A.7}$$

where f_{c} is defined in Eq. (20) and k_i are expressed as

$$k_1 = 1 + \frac{\partial\eta^{\text{ir}}}{\partial\eta^{\text{rev}}}, \quad k_2 = \eta^{\text{ir}} + \eta^{\text{rev}}\frac{\partial\eta^{\text{ir}}}{\partial\eta^{\text{rev}}}, \quad k_3 = \eta^{\text{ir}} + \frac{1}{2}\eta^{\text{rev}}\frac{\partial\eta^{\text{ir}}}{\partial\eta^{\text{rev}}}, \quad k_4 = \epsilon_{\text{T}} + \frac{\partial\varepsilon^{\text{p}}}{\partial\eta^{\text{rev}}}. \tag{A.8}$$

It is important to highlight that the necessary condition for the minimum of π with respect to η^{rev}, which leads to the transformation criteria (A.7), is not computed in a standard manner. This stems from the state-dependence of the dissipation potential D, i.e., the dependence of f_{c} on the accumulated volume fraction η^{acc}, see Eq. (20). Having the minimization problem formulated in rates (not shown here), it becomes apparent that f_{c} is treated as a constant when evaluating the necessary condition for the rate $\dot{\eta}^{\text{rev}}$. In the incremental setting, to maintain consistency with the rate-problem, the increment of the martensite volume fraction, $\Delta\eta^{\text{rev}}$, present in the current unknown $\eta^{\text{rev}} = \Delta\eta^{\text{rev}} + \eta^{\text{rev}}_n$ is distinguished from the increment upon which the evolution equation for f_{c} rely. Despite the two increments coincide, the latter is considered as constant when evaluating the necessary condition. Accordingly, the minimization problem does posses the structure of a quasi-optimization problem and not a genuine optimization problem. To avoid the complexity in the model presentation, this issue is not elaborated here. It should be remarked that upon assuming the same increment $\Delta\eta^{\text{rev}}$ for the current unknown η^{rev} and f_{c}, resulting in a non-constant f_{c} in the calculation of the necessary condition, extra differentiation terms arise in the transformation criteria (A.7). However, our auxiliary simulations showed that these extra terms only marginally contribute to the results.

References

[1] K. Bhattacharya. *Microstructure of martensite: why it forms and how it gives rise to the shape-memory effect.* Oxford University Press, Oxford, 2003.

[2] J. Mohd Jani, M. Leary, A. Subic, and M. A. Gibson. A review of shape memory alloy research, applications and opportunities. *Mat. Des.*, 56:1078–1113, 2014.

[3] K. Nargatti and S. Ahankari. Advances in enhancing structural and functional fatigue resistance of superelastic NiTi shape memory alloy: A review. *J. Intell. Mat. Sys. Struct.*, 33: 503–531, 2022.

[4] G. Eggeler, E. Hornbogen, A. Yawny, A. Heckmann, and M. Wagner. Structural and functional fatigue of NiTi shape memory alloys. *Mat. Sci. Engng. A*, 378:24–33, 2004.

[5] A. R. Pelton. Nitinol fatigue: A review of microstructures and mechanisms. *J. Mater. Eng. Perform.*, 20:613–617, 2011.

[6] M. J. Mahtabi, N. Shamsaei, and M. R. Mitchell. Fatigue of Nitinol: The state-of-the-art and ongoing challenges. *J. Mech. Behav. Biomed. Mater.*, 50:228–254, 2015.

[7] G. Kang and D. Song. Review on structural fatigue of NiTi shape memory alloys: Pure mechanical and thermo-mechanical ones. *Theor. Appl. Mech. Lett.*, 5:245–254, 2015.

[8] L. Zheng, Y. He, and Z. Moumni. Investigation on fatigue behaviors of NiTi polycrystalline strips under stress-controlled tension via in-situ macro-band observation. *Int. J. Plast.*, 90: 116–145, 2017.

[9] R. Sidharth, A. S. K. Mohammed, and H. Sehitoglu. Functional fatigue of NiTi shape memory alloy: effect of loading frequency and source of residual strains. *Shap. Mem. Superelasticity*, 8: 394–412, 2022.

[10] Y. Zhang, X. Chai, X. Ju, Y. You, S. Zhang, L. Zheng, Z. Moumni, J. Zhu, and W. Zhang. Concentration of transformation-induced plasticity in pseudoelastic NiTi shape memory alloys: Insight from austenite–martensite interface instability. *Int. J. Plast.*, 160:103481, 2023.

[11] Q. Kan, Y. Zhang, Y. Xu, G. Kang, and C. Yu. Tension-compression asymmetric functional degeneration of super-elastic NiTi shape memory alloy: Experimental observation and multiscale constitutive model. *Int. J. Solids Struct.*, 280:112384, 2023.

[12] H. Akamine, A. Heima, Y. Soejima, M. Mitsuhara, T. Inamura, and M. Nishida. Where and when are dislocations induced by thermal cycling in Ti–Ni shape memory alloys? *Acta Mater.*, 244:118588, 2023.

[13] R. Sidharth, T. B. Celebi, and H. Sehitoglu. Origins of functional fatigue and reversible transformation of precipitates in NiTi shape memory alloy. *Acta Mater.*, page 119990, 2024.

[14] J. A. Shaw and S. Kyriakides. On the nucleation and propagation of phase transformation fronts in a NiTi alloy. *Acta Mater.*, 45:683–700, 1997.

[15] J. A. Shaw and S. Kyriakides. Initiation and propagation of localized deformation in elasto-plastic strips under uniaxial tension. *Int. J. Plast.*, 13:837–871, 1997.

[16] Q. P. Sun and Z. Q. Li. Phase transformation in superelastic NiTi polycrystalline micro-tubes under tension and torsion—from localization to homogeneous deformation. *Int. J. Solids Struct.*, 39:3797–3809, 2002.

[17] B. Reedlunn, C. B. Churchill, E. E. Nelson, J. A. Shaw, and S. H. Daly. Tension, compression, and bending of superelastic shape memory alloy tubes. *J. Mech. Phys. Solids*, 63:506–537, 2014.

[18] P. H. Lin, H. Tobushi, K. Tanaka, T. Hattori, and M. Makita. Pseudoelastic behaviour of TiNi shape memory alloy subjected to strain variations. *J. Intell. Mat. Sys. Struct.*, 5:694–701, 1994.

[19] T. J. Lim and D. L. McDowell. Degradation of an Ni-Ti alloy during cyclic loading. In Vijay K. Varadan, editor, *Smart Structures and Materials 1994: Smart Materials*, volume 2189, pages 326 – 341. International Society for Optics and Photonics, SPIE, 1994.

[20] M. A. Iadicola and J. A. Shaw. The effect of uniaxial cyclic deformation on the evolution of phase transformation fronts in pseudoelastic NiTi wire. *J. Intell. Mat. Sys. Struct.*, 13:143–155, 2002.

[21] L. C. Brinson, I. Schmidt, and R. Lammering. Stress-induced transformation behavior of a polycrystalline NiTi shape memory alloy: micro and macromechanical investigations via in situ optical microscopy. *Int. J. Solids Struct.*, 52:1549–1571, 2004.

[22] L. Zheng, Y. He, and Z. Moumni. Effects of Lüders-like bands on NiTi fatigue behaviors. *Int. J. Solids Struct.*, 83:28–44, 2016.

[23] L. Zheng, Y. He, and Z. Moumni. Lüders-like band front motion and fatigue life of pseudoelastic polycrystalline NiTi shape memory alloy. *Scripta Mat.*, 123:46–50, 2016.

[24] H. Tobushi, T. Ikawa, and R. Matsui. Pseudoviscoelastic behavior of TiNi shape memory alloy in subloop. *Trans. Mater. Res. Soc. Jpn.*, 28:611–614, 2003.

[25] S. Doraiswamy, A. Rao, and A. R. Srinivasa. Combining thermodynamic principles with Preisach models for superelastic shape memory alloy wires. *Smart Mater. Struct.*, 20:085032, 2011.

[26] K. Takeda, H. Tobushi, K. Miyamoto, and E. A. Pieczyska. Superelastic deformation of TiNi shape memory alloy subjected to various subloop loadings. *Mater. Trans.*, 53:217–223, 2012.

[27] J. Ortin. Partial hysteresis cycles in shape-memory alloys: Experiments and modelling. *J. Phys. IV France*, 1(C4):65–70, 1991.

[28] I. Müller and H. Xu. On the pseudo-elastic hysteresis. *Acta Metall. Mater.*, 39:263–271, 1991.

[29] M. A. Savi and A. Paiva. Describing internal subloops due to incomplete phase transformations in shape memory alloys. *Arch. Appl. Mech.*, 74:637–647, 2005.

[30] L. Saint-Sulpice, S. Arbab Chirani, and S. Calloch. A 3D super-elastic model for shape memory alloys taking into account progressive strain under cyclic loadings. *Mech. Mat.*, 41:12–26, 2009.

[31] C. W. Chan, S. H. J. Chan, H. C. Man, and P. Ji. 1-D constitutive model for evolution of stress-induced R-phase and localized Lüders-like stress-induced martensitic transformation of super-elastic NiTi wires. *Int. J. Plast.*, 32:85–105, 2012.

[32] T. Bartel, M. Osman, and A. Menzel. A phenomenological model for the simulation of functional fatigue in shape memory alloy wires. *Meccanica*, 52:973–988, 2017.

[33] C. Bouvet, S. Calloch, and C. Lexcellent. A phenomenological model for pseudoelasticity of shape memory alloys under multiaxial proportional and nonproportional loadings. *Eur. J. Mech. A/Solids*, 23:37–61, 2004.

[34] S. Stupkiewicz and H. Petryk. A bi-crystal aggregate model of pseudoelastic behaviour of shape-memory alloy polycrystals. *Int. J. Mech. Sci.*, 52:219–228, 2010.

[35] J. F. Hallai and S. Kyriakides. Underlying material response for Lüders-like instabilities. *Int. J. Plast.*, 47:1–12, 2013.

[36] J. L. Greenly, S. Kyriakides, and S. Tsimpoukis. On the underlying material response of pseudoelastic NiTi. *Eur. J. Mech. A/Solids*, 104:105023, 2024.

[37] Y. Xiao and D. Jiang. Constitutive modelling of transformation pattern in superelastic NiTi shape memory alloy under cyclic loading. *Int. J. Mech. Sci.*, 182:105743, 2020.

[38] Y. Xiao and D. Jiang. Thermomechanical modeling on cyclic deformation and localization of superelastic NiTi shape memory alloy. *Int. J. Solids Struct.*, 250:111723, 2022.

[39] S. Stupkiewicz and H. Petryk. A robust model of pseudoelasticity in shape memory alloys. *Int. J. Num. Meth. Engng.*, 93:747–769, 2013.

[40] M. Rezaee Hajidehi and S. Stupkiewicz. Gradient-enhanced model and its micromorphic regularization for simulation of Lüders-like bands in shape memory alloys. *Int. J. Solids Struct.*, 135:208–218, 2018.

[41] M. Rezaee-Hajidehi, K. Tůma, and S. Stupkiewicz. Gradient-enhanced thermomechanical 3D model for simulation of transformation patterns in pseudoelastic shape memory alloys. *Int. J. Plast.*, 128:102589, 2020.

[42] M. Rezaee-Hajidehi and S. Stupkiewicz. Modelling of propagating instabilities in pseudoelastic NiTi tubes under combined tension–torsion: helical bands and apparent yield locus. *Int. J. Solids Struct.*, 221:130–149, 2021.

[43] M. Rezaee-Hajidehi and S. Stupkiewicz. Predicting transformation patterns in pseudoelastic NiTi tubes under proportional axial–torsion loading. *Int. J. Solids Struct.*, 281:112436, 2023.

[44] R. Delville, B. Malard, J. Pilch, P. Sittner, and D. Schryvers. Transmission electron microscopy investigation of dislocation slip during superelastic cycling of Ni–Ti wires. *Int. J. Plast.*, 27: 282–297, 2011.

[45] Y. Zhang, W. Li, Z. Moumni, J. Zhu, W. Zhang, and S. Y. Zhong. Degradation of the recoverable strain during stress controlled full transformation cycling in NiTi shape memory alloys. *Scripta Mat.*, 162:68–71, 2019.

[46] P. Sedmák, P. Šittner, J. Pilch, and C. Curfs. Instability of cyclic superelastic deformation of NiTi investigated by synchrotron X-ray diffraction. *Acta Mater.*, 94:257–270, 2015.

[47] X. Wang, B. Xu, and Z. Yue. Phase transformation behavior of pseudoelastic NiTi shape memory alloys under large strain. *J. Alloys Compd.*, 463:417–422, 2008.

[48] G. Kang, Q. Kan, L. Qian, and Y. Liu. Ratchetting deformation of super-elastic and shape-memory NiTi alloys. *Mech. Mat.*, 41:139–153, 2009.

[49] W. Zaki and Z. Moumni. A 3D model of the cyclic thermomechanical behavior of shape memory alloys. *J. Mech. Phys. Solids*, 55:2427–2454, 2007.

[50] Q. Kan and G. Kang. Constitutive model for uniaxial transformation ratchetting of super-elastic NiTi shape memory alloy at room temperature. *Int. J. Plast.*, 26:441–465, 2010.

[51] G. Scalet, A. Karakalas, L. Xu, and D. Lagoudas. Finite strain constitutive modelling of shape memory alloys considering partial phase transformation with transformation-induced plasticity. *Smart Mater. Struct.*, 7:206–221, 2021.

[52] A. Sadjadpour and K. Bhattacharya. A micromechanics-inspired constitutive model for shape-memory alloys. *Smart Mater. Struct.*, 16:1751, 2007.

[53] F. Auricchio, A. Reali, and U. Stefanelli. A three-dimensional model describing stress-induced solid phase transformation with permanent inelasticity. *Int. J. Plast.*, 23:207–226, 2007.

[54] L. Petrini and A. Bertini. A three-dimensional phenomenological model describing cyclic behavior of shape memory alloys. *Int. J. Plast.*, 125:348–373, 2020.

[55] H. Petryk. Incremental energy minimization in dissipative solids. *C. R. Mecanique*, 331:469–474, 2003.

[56] S. Forest. Micromorphic approach for gradient elasticity, viscoplasticity, and damage. *J. Eng. Mech.*, 135:117–131, 2009.

[57] M. Mazière and S. Forest. Strain gradient plasticity modeling and finite element simulation of Lüders band formation and propagation. *Continuum Mech. Thermodyn.*, 27:83–104, 2015.

[58] J. Korelc. Automation of primal and sensitivity analysis of transient coupled problems. *Comp. Mech.*, 44:631–649, 2009.

[59] J. Korelc and P. Wriggers. *Automation of Finite Element Methods*. Springer International Publishing, Switzerland, 2016.

[60] P. Sedmák, J. Pilch, L. Heller, J. Kopeček, J. Wright, P. Sedlák, M. Frost, and P. Šittner. Grain-resolved analysis of localized deformation in nickel-titanium wire under tensile load. *Science*, 353:559–562, 2016.

[61] S. Stupkiewicz, M. Rezaee-Hajidehi, and H. Petryk. Multiscale analysis of the effect of interfacial energy on non-monotonic stress–strain response in shape memory alloys. *Int. J. Solids Struct.*, 221:77–91, 2021.

[62] C. Morin, Z. Moumni, and W. Zaki. Thermomechanical coupling in shape memory alloys under cyclic loadings: Experimental analysis and constitutive modeling. *Int. J. Plast.*, 27:1959–1980, 2011.

[63] Q. Kan, C. Yu, G. Kang, J. Li, and W. Yan. Experimental observations on rate-dependent cyclic deformation of super-elastic NiTi shape memory alloy. *Mech. Mat.*, 97:48–58, 2016.

[64] P. Šittner, P. Sedlák, H. Seiner, P. Sedmák, J. Pilch, R. Delville, L. Heller, and L. Kadeřávek. On the coupling between martensitic transformation and plasticity in NiTi: Experiments and continuum based modelling. *Prog. Mat. Sci.*, 98:249–298, 2018.

[65] R. T. Watkins, B. Reedlunn, S. Daly, and J. A. Shaw. Uniaxial, pure bending, and column buckling experiments on superelastic NiTi rods and tubes. *Int. J. Solids Struct.*, 146:1–28, 2018.

[66] P. Sedlák, M. Frost, M. Ševčík, and H. Seiner. 3D spatial reconstruction of macroscopic austenite–martensite transition zones in NiTi wires induced by tension and twisting using diffraction/scattering computed tomography. *Int. J. Solids Struct.*, 228:111122, 2021.

[67] R. Hill. On discontinuous plastic states, with special reference to localized necking in thin sheets. *J. Mech. Phys. Solids*, 1:19–30, 1952.

[68] X. Zhang, P. Feng, Y. He, T. Yu, and Q. Sun. Experimental study on rate dependence of macroscopic domain and stress hysteresis in NiTi shape memory alloy strips. *Int. J. Mech. Sci.*, 52:1660–1670, 2010.

[69] N. J. Bechle and S. Kyriakides. Localization in NiTi tubes under bending. *Int. J. Mech. Sci.*, 51:967–980, 2014.

[70] H. Ahmadian, S. Hatefi Ardakani, and S. Mohammadi. Strain-rate sensitivity of unstable localized phase transformation phenomenon in shape memory alloys using a non-local model. *Int. J. Solids Struct.*, 63:167–183, 2015.

[71] D. Jiang, S. Kyriakides, C. M. Landis, and K. Kazinakis. Modeling of propagation of phase transformation fronts in NiTi under uniaxial tension. *Eur. J. Mech. A/Solids*, 64:131–142, 2017.