Phase field modelling of fracture and fatigue in Shape Memory Alloys

Marlini Simoes\textsuperscript{a}, Emilio Martínez-Pañeda\textsuperscript{b,}\textsuperscript{*}

\textsuperscript{a} Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE, UK
\textsuperscript{b} Department of Civil and Environmental Engineering, Imperial College London, London SW7 2AZ, UK

Received 8 August 2020; received in revised form 5 October 2020; accepted 9 October 2020

Abstract

We present a new phase field framework for modelling fracture and fatigue in Shape Memory Alloys (SMAs). The constitutive model captures the superelastic behaviour of SMAs and damage is driven by the elastic and transformation strain energy densities. We consider both the assumption of a constant fracture energy and the case of a fracture energy dependent on the martensitic volume fraction. The framework is implemented in an implicit time integration scheme, with both monolithic and staggered solution strategies. The potential of this formulation is showcased by modelling a number of paradigmatic problems. First, a boundary layer model is used to examine crack tip fields and compute crack growth resistance curves (R-curves). We show that the model is able to capture the main fracture features associated with SMAs, including the toughening effect associated with stress-induced phase transformation. Insight is gained into the role of temperature, material strength, crack density function and fracture energy homogenisation. Secondly, several 2D and 3D boundary value problems are addressed, demonstrating the capabilities of the model in capturing complex cracking phenomena in SMAs, such as unstable crack growth, mixed-mode fracture or the interaction between several cracks. Finally, the model is extended to fatigue and used to capture crack nucleation and propagation in biomedical stents, a paradigmatic application of nitinol SMAs.

© 2020 Elsevier B.V. All rights reserved.

Keywords: Phase field; Shape Memory Alloys; Fracture; Fatigue; Finite element analysis

1. Introduction

Shape Memory Alloys (SMAs) have gained increasing attention in recent years, with applications spanning the areas of aerospace, bioengineering, transport and infrastructure, among others [1]. The popularity of these smart, multi-functional materials is largely grounded on their capacity to sustain notably large recoverable strains (up to 10\%) as a result of transformation between their austenitic and martensitic phases. This transformation can be attained by changing the mechanical load (stress-induced transformation), the temperature (temperature-induced transformation) or both (stress and temperature-induced transformation). As shown in Fig. 1a, two phenomena are intrinsic to SMAs and their phase transformation: superelasticity (SE) and shape memory effect (SME). In both cases, due to the creation of a stress-induced phase, the application of a mechanical load renders a non-linear response with very large strains. However, only superelastic alloys, such as nickel–titanium (nitinol), can achieve full strain recovery when the load is removed, exhibiting a hysteresis loop in the stress versus strain response.

\textsuperscript{*} Corresponding author.

E-mail address: e.martinez-paneda@imperial.ac.uk (E. Martínez-Pañeda).

https://doi.org/10.1016/j.cma.2020.113504
0045-7825/© 2020 Elsevier B.V. All rights reserved.
SMAs experiencing the shape memory effect show a large residual strain after unloading and require a change of temperature to recover their original shape.

In SMAs where both superelasticity and shape memory effect can take place, their occurrence is governed by the temperature of the system, $T$. Fig. 1b shows a typical SMA stress–temperature curve, where $M_f$, $M_s$, $A_s$ and $A_f$ denote the martensite end and start temperatures, and the austenite start and end temperatures, respectively. If $T > A_f$ (point 1), superelasticity will be observed; upon loading, the material will be in the austenite phase until reaching the stress level dictating the start of the transformation $\sigma_{st L}$ (point 2), and the transformation will end upon attaining $\sigma_{st U}$ (point 3), when the material will be fully martensitic. If the load is removed, full recovery is observed; reverse transformation starts at $\sigma_{st L}$ (point 4) and all the martensite transforms into austenite upon reaching $\sigma_{st U}$ (point 5). The shape memory effect will be observed if $T < A_s$; austenite to martensite transformation will start when the stress level reaches the threshold $\sigma_{st L}$ (point B), ending upon attaining $\sigma_{st U}$ (point C). However, no reverse transformation takes place upon unloading since austenite is not stable at this temperature; $T$ must be raised to eliminate the residual strains. The magnitude of the temperature-dependent stress thresholds ($\sigma_{st L}$, $\sigma_{st U}$, $\sigma_{st L}$, $\sigma_{st U}$) is governed by $C_M$ and $C_A$, the slope of the stress–temperature diagram for martensite and austenite, respectively. The material behaviour in the phase transformation region is typically defined as a function of the martensitic volume fraction, $\xi$ (0 $\leq \xi \leq 1$). The reader is referred to Refs. [2–5] for more details.

The fracture and fatigue behaviour of SMAs have attracted significant interest, from both numerical and experimental perspectives; see, for example, the reviews by Robertson et al. [6] and Baxevanis and Lagoudas [7]. Since the yield stress of SMAs is typically much larger than the transformation stress $\sigma_{fL}$ [8], a stress-induced transformation zone develops in the vicinity of the crack tip. This crack tip transformation region has been characterised using infrared (IR) thermography [9] and synchrotron X-ray diffraction [10,11]. The high stresses of the transformation region dominate the crack growth resistance of SMAs and result in energy dissipation and material toughening [12,13]. Finite element models have been developed to predict the role of transformation toughening and reverse transformation on crack propagation [5,14–16]. While important insight has been gained, these efforts have been restricted to discrete numerical methods, such as cohesive zone formulations. Discrete numerical methods for fracture are limited when dealing with the complex conditions of practical applications and, consequently, important challenges remain unaddressed (crack nucleation, mixed-mode, interacting cracks, etc.). Moreover, to the best of the authors’ knowledge, a modelling framework capable of explicitly predicting fatigue crack growth behaviour in SMAs has not been presented yet. The main objective of this study is to address these two important knowledge gaps by developing a phase field-based computational framework for fracture and fatigue cracking in SMAs.

The phase field method has proven to be a compelling variational framework for predicting advanced fracture problems. The classical Griffith fracture energy balance [17] is revisited as an energy minimisation problem by
solving for an auxiliary variable, the phase field parameter $\phi$ [18,19]. This enables capturing, on the original finite element mesh, complex cracking phenomena such as crack nucleation, branching, kinking or merging in arbitrary geometries and dimensions (see, e.g., [20–23]). Not surprisingly, the method is enjoying great popularity and its success has been extended to numerous applications. Recent examples include fracture of functionally graded materials [24,25], composites delamination [26,27], cracking in solar-grade silicon [28], hydrogen embrittlement [29–31], rock fracture [32,33], and fatigue damage [34,35], among others; see [36] for a review.

In this work, we present the first phase field formulation for fracture (and fatigue) in Shape Memory Alloys (SMAs). The constitutive behaviour of the solid includes both stress and temperature-induced phase transformations, capturing the superelasticity (reverse transformation) and shape memory effects. The evolution of the phase field variable is driven by the combination of elastic and transformation strain energy densities. Two different phase field dissipation functions are considered, corresponding to the so-called AT1 [37] and AT2 [19,38] models, and their influence is assessed. The model is implemented in an implicit time integration scheme and the coupled displacement–phase field problem is solved using both monolithic (quasi-Newton) and staggered schemes, demonstrating the robustness of the framework. Several paradigmatic boundary value problems are addressed to demonstrate the potential of the framework in providing physical insight into the fracture behaviour of SMAs and covering complex, large scale 3D fatigue problems. The remainder of the paper is organised as follows. The theoretical framework is presented in Section 2. Then, the finite element implementation is described in Section 3. Representative numerical results are shown in Section 4. First, a boundary layer model is used to gain insight into the fracture problem of a biomedical stent. Concluding remarks end the paper in Section 5.

2. A phase field fracture formulation for Shape Memory Alloys

2.1. Constitutive behaviour of SMAs

To constitutively describe the material behaviour of SMAs we follow the so-called unified model by Lagoudas and co-workers [1,39,40], including recent extensions to capture gradual phase transformations, and the stress-dependencies of the inelastic recoverable strain and the phase diagram slope [41,42]. The model builds upon the rule of mixtures to determine the magnitude of relevant properties for material points located in the phase transformation region. Thus, the effective value of any relevant phase-dependent parameter ($\Theta$) is a function of the martensitic volume fraction $\xi$ (0 \leq \xi \leq 1) and its magnitude in the austenitic ($\Theta^A$) and martensitic ($\Theta^M$) phases:

$$\Theta(\xi) = (1 - \xi) \Theta^A + \xi \Theta^M. \quad (1)$$

2.1.1. Thermodynamic potential

A Gibbs free energy potential $G$ can be defined as a function of the terms corresponding to the austenitic ($G^A$) and martensitic ($G^M$) phases, the martensitic volume fraction ($\xi$), and a mixing term due to the phase transformation ($G^{mix}$). Both $G^A$ and $G^M$ are functions of the Cauchy stress tensor $\sigma$ and the absolute temperature $T$, while $G^{mix}$ is a function of $\sigma$, the transformation strain tensor $\varepsilon'$, and the so-called transformation hardening energy $g'$, which is a measure of the nonlinear change in the mixing energy as transformation progresses at constant stress [42]. Thus, the Gibbs energy reads,

$$G(\sigma, T, \varepsilon', \xi, g') = (1 - \xi) G^A(\sigma, T) + \xi G^M(\sigma, T) + G^{mix}(\sigma, \varepsilon', g'). \quad (2)$$

Assuming a quadratic dependence on stress [42], for a material with density $\rho$, thermal expansion tensor $\alpha$, compliance tensor $S$, and specific heat $c$, the $G^A (\gamma = A)$ and $G^M (\gamma = M)$ terms read,

$$G^\gamma(\sigma, T) = -\frac{1}{2\rho} \sigma : S^\gamma - \frac{1}{\rho} \sigma : \alpha (T - T_0) + c^\gamma \left[ (T - T_0) - T \ln \left( \frac{T}{T_0} \right) \right] - s^{\gamma}_0 T + u^{\gamma}_0, \quad (3)$$

where $s_0$, $u_0$, and $T_0$ respectively denote the specific entropy, specific internal energy and temperature at the reference state. Finally, the mixing term of the Gibbs free energy is given by,

$$G^{mix}(\sigma, \varepsilon', g') = -\frac{1}{\rho} \sigma : \varepsilon' + \frac{g'}{\rho}. \quad (4)$$
2.1.2. Evolution equations

Evolution equations for the transformation strain \( \varepsilon^t \) and the hardening variable \( g^t \) are now provided. \( \varepsilon^t \), an inelastic strain tensor generated during transformation from austenite to martensite, is a function of the rate of the martensitic volume fraction \( \dot{\xi} \) and the so-called transformation direction tensor \( \Lambda^t \):

\[
\dot{\varepsilon}^t = \Lambda^t \dot{\xi}.
\]  

If the rate of the martensitic volume fraction is positive (\( \dot{\xi} > 0 \)), forward transformation takes place (\( \Lambda^t = \Lambda^t_{fwd} \)) and the transformation direction tensor equals

\[
\Lambda^t_{fwd} = \frac{3}{2} H^{cur} \frac{\sigma^t}{\sigma_e},
\]  

where the prime symbol \( ' \) denotes deviatoric quantities, \( \sigma_e \) is an effective stress, defined as in von Mises plasticity \( \sigma_e = \sqrt{\frac{3}{2} \sigma' : \sigma'} \), and \( H^{cur} \) is the uniaxial transformation strain magnitude for complete transformation. On the other hand, if \( \dot{\xi} < 0 \), then the transformation direction tensor corresponds to its reverse form (\( \Lambda^t = \Lambda^t_{rev} \)), defined as a function of the transformation strain and the martensite volume fraction at the reversal:

\[
\Lambda^t_{rev} = \frac{\varepsilon^t - \xi}{\xi}. \tag{5}
\]

The modelling framework has the capability of capturing the stress sensitivity of the maximum transformation strain. This is achieved by defining \( H^{cur} \) as a decaying exponential function when the effective stress exceeds a critical quantity \( \sigma_c \) [41]. Thus, \( H^{cur} \) is given by:

\[
H^{cur} (\bar{\sigma}) = \begin{cases} H_{min} & \text{if } \sigma_e \leq \sigma_c, \\ H_{min} + \left[ H_{max} - H_{min} \left(1 - \exp (k(\sigma_e - k\sigma_c)) \right) \right] & \text{if } \sigma_e > \sigma_c, \end{cases} \tag{6}
\]

where \( H_{min} \) corresponds to the observable uniaxial two-way shape memory effect (TWSME), \( k \) is a fitting parameter and \( H_{max} \) is the ultimate transformation strain under uniaxial loading.

It remains to define an evolution equation for the transformation hardening energy variable \( g^t \); this is achieved by means of a hardening function \( f^t \), which takes distinct values during forward and reverse transformation. Thus,

\[
\dot{g}^t = f^t \dot{\xi}, \tag{7}
\]

with the hardening function being of the form,

\[
f^t_{fwd} (\xi) = \frac{a_1}{2} \left(1 + \xi^{n_1} - (1 - \xi)^{n_2}\right) + a_3
\]

where \( \dot{\xi} > 0 \) (\( f^t = f^t_{fwd} \)); while for reverse transformation (\( \dot{\xi} < 0 \), \( f^t = f^t_{rev} \)) the hardening function reads:

\[
f^t_{rev} (\xi) = \frac{a_2}{2} \left(1 + \xi^{n_3} - (1 - \xi)^{n_4}\right) - a_3. \tag{8}
\]

Here, the exponents \( n_1, n_2, n_3 \) and \( n_4 \) take real numbers in the range \((0, 1]\) and are aimed at enabling the simulation of gradual hardening behaviour during transformation. The richer description provided, relative to other existing models (exponential, trigonometric, linear, etc.), enables matching the experimental data more closely [41]. Apart from that, the constants \( a_1-a_3 \) are computed from the common SMA material parameters \( M_s, M_f, A_s, A_f, C_A, \) and \( C_M \), as described in Ref. [41].

2.1.3. Thermodynamically-consistent constitutive prescriptions

A suitable free energy can be defined building upon a thermodynamically-consistent energy imbalance, as elaborated elsewhere [42]. Accordingly, the total infinitesimal strain can be defined as follows:

\[
\varepsilon = -\rho \partial_T \mathcal{G} = \mathbf{S} : \sigma + \alpha (T - T_0) + \varepsilon', \tag{9}
\]

incorporating the contributions from the elastic, thermal and transformation strains. And the constitutive relation for the entropy reads,

\[
s = -\partial_T \mathcal{G} = \frac{1}{\rho} \alpha : \sigma + c \ln \left(\frac{T}{T_0}\right) + s_0. \tag{10}
\]
The criteria determining the onset of transformation are given by,

\[
\dot{\xi} \phi = 0 , \quad \text{with} \quad \phi = \begin{cases} 
\Phi_{fwd} = \pi' - Y \leq 0 & \text{if } \dot{\xi} > 0 \\
\Phi_{rev} = -\pi' - Y \leq 0 & \text{if } \dot{\xi} < 0 ,
\end{cases}
\]

where \( \Phi \) is the transformation surface and \( \pi' \) is the thermodynamic driving force for transformation, work conjugate to \( \xi \). The latter can be defined as:

\[
\pi' = \mathbf{\sigma} : \mathbf{\Lambda}' + \frac{1}{2} \mathbf{\sigma} : (\mathbf{S}^M - \mathbf{S}^A) : \mathbf{\sigma} + \rho \left( s_0^M - s_0^A \right) T - \rho \left( u_0^M - u_0^A \right) - f',
\]

where the transformation direction tensor \( \mathbf{\Lambda}' \) and the hardening function \( f' \) take their forward or reverse form for \( \dot{\xi} > 0 \) and \( \dot{\xi} < 0 \), respectively.

2.2. Variational phase field fracture

Fracture and fatigue in SMAs are predicted by using a variational phase field model [19,43,44]. Since the early work by Francfort and Marigo [18], phase field fracture models aim at providing a variational framework for the concept of crack advance driven by the competition between toughness (surface energy density, \( G \)) and energy release rate \( \dot{G} \); as first proposed by Griffith [17] for elastic solids, and later extended to account for inelastic energy dissipation by Orowan [45]. In the present study, we will consider both a constant material toughness \( G_c \) and the case in which the critical energy release rate is determined from its austenite and martensite counterparts, \( G_c(\xi) \), using the rule of mixtures.

The discrete crack is approximated through an auxiliary field variable \( \phi \), which varies between \( \phi = 0 \), intact material, and \( \phi = 1 \), fully cracked material. The size of the regularised crack surface is governed by the choice of \( \ell \), the phase field model-inherent length scale. Thus, the fracture energy due to the creation of a crack can be approximated as:

\[
\int_{\Gamma} G_c(\xi) \, d\Gamma \approx \int_{\Omega} G_c(\xi) \gamma(\phi, \nabla\phi) \, dV = \int_{\Omega} \frac{G_c(\xi)}{4c_w} \left( \frac{w(\phi)}{\ell} + \ell|\nabla\phi|^2 \right) \, dV ,
\]

where \( \gamma(\phi, \nabla\phi) \) is the crack density function, which is itself a function of \( w(\phi) \) and \( c_w \). The most exploited constitutive choices of \( w(\phi) \) and \( c_w \) are those associated with the so-called AT1 [37] and AT2 [19,38] models:

\[
\begin{align*}
\text{AT1:} \quad & w(\phi) = \phi , \quad c_w = 2/3 \\
\text{AT2:} \quad & w(\phi) = \phi^2 , \quad c_w = 1/2
\end{align*}
\]

The main difference between them is the fact that the AT1 model has a non-zero elastic limit. Other constitutive choices have been proposed, such as the so-called phase-field regularised cohesive zone model (PF-CZM) [46]; the reader is referred to Ref. [47] for a detailed numerical comparison. To retain generality, we proceed to present our phase field formulation without making specific constitutive choices for \( w(\phi) \) and \( c_w \), as both the AT1 and the AT2 models will be considered in this study.

The total potential energy can be expressed as a function of the contributions from the mechanical, thermal and fracture terms as:

\[
\Psi = \Psi^s(\mathbf{e}, \xi, \phi) + \Psi^T(T, \xi) + \Psi^\phi(\phi, \xi) = \int_{\Omega} \left\{ (1 - \phi)^2 \Psi(\mathbf{e}, T, \xi) + c(\xi) \left[ (T - T_0) - T \ln \left( \frac{T}{T_0} \right) \right] + \frac{G_c(\xi)}{4c_w} \left( \frac{w(\phi)}{\ell} + \ell|\nabla\phi|^2 \right) \right\} \, dV ,
\]

where \( \Psi(\mathbf{e}, T, \xi) \) is the strain energy density of the solid. Recall, see Eq. (12), that the strain tensor additively decomposes into an elastic part \( \mathbf{e}^{es} \), a thermal part \( \mathbf{e}^T = \alpha \Delta T \) and a transformation part \( \mathbf{e}' \). Accordingly, the total strain energy density \( \Psi \) can be expressed as

\[
\Psi(\mathbf{e}, T, \xi) = \int_0^t (\mathbf{e} : \dot{\mathbf{e}}^e) \, dt + \int_0^t (\mathbf{e} : \dot{\mathbf{e}}^T) \, dt + \int_0^t (\mathbf{e} : \dot{\mathbf{e}}') \, dt .
\]

As elaborated below, in this work we assume that fracture is driven by the total strain energy density.
Consider now the total potential energy of the solid, Eq. (19). The strong form can be readily derived by taking the first variation with respect to $\varepsilon, T$ and $\phi$, and making use of Gauss’ divergence theorem. Thus, the coupled field equations read,

\begin{align*}
(1 - \phi)^2 \nabla \cdot \sigma &= 0 \quad \text{in} \quad \Omega \\
\rho c (\xi) \dot{T} + \nabla \cdot q &= 0 \quad \text{in} \quad \Omega \\
G_c (\xi) \left( \frac{\phi}{\ell} - \ell \Delta \phi \right) - 2(1 - \phi) \psi &= 0 \quad \text{in} \quad \Omega
\end{align*}

(21)

where $q$ is the heat flux per unit area of the solid.

2.2.1. Phase field fatigue

Now we proceed to incorporate a fatigue degradation function $f(\bar{\alpha}(t))$, extending the recent work by Carrara et al. [35] to SMAs. The field equation corresponding to the phase field variable, (21c) is therefore given by:

\begin{equation}
\frac{f(\bar{\alpha}(t)) G_c (\xi) \left( \frac{\phi}{\ell} - \ell \Delta \phi \right)}{\bar{\alpha}(t)} - 2(1 - \phi) \psi = 0,
\end{equation}

(22)

where $\bar{\alpha}$ is a cumulation of any scalar quantity which can describe the fatigue history experienced by the material. For a pseudo-time $\tau$, the cumulative history variable can be defined as [35,48]:

\begin{equation}
\bar{\alpha}(t) = \int_0^t H(\alpha \dot{\alpha}) \dot{\alpha} |\dot{\alpha}| \, d\tau,
\end{equation}

(23)

where $H(\alpha \dot{\alpha})$ is the Heaviside step function, defined as $H(\alpha \dot{\alpha}) = 1$ if $\alpha \dot{\alpha} \geq 0$ (loading) and $H(\alpha \dot{\alpha}) = 0$ otherwise (unloading). Accordingly, $\bar{\alpha}$ only grows during loading. It remains to define a fatigue history variable $\alpha$, representing the loading condition in the solid, and a fatigue degradation function $f(\bar{\alpha}(t))$, characterising the sensitivity of the fracture energy to the number of cycles. Regarding the former, we follow an energetic approach and define $\alpha = g(\phi) \psi$. While the degradation function $f(\bar{\alpha}(t))$ is chosen so as to vanish asymptotically,

\begin{equation}
f(\bar{\alpha}(t)) = \begin{cases} 
1 & \text{if} \quad \bar{\alpha}(t) \leq \alpha_T \\
\left( \frac{2 \alpha_T}{\bar{\alpha}(t) + \alpha_T} \right)^2 & \text{if} \quad \bar{\alpha}(t) \leq \alpha_T.
\end{cases}
\end{equation}

(24)

Here, $\alpha_T$ represents a threshold value, below which the fracture energy remains unaffected; as in Ref. [23], we define it as:

\begin{equation}
\alpha_T = \frac{G_c}{12 \ell}
\end{equation}

(25)

3. Finite element implementation

We proceed to describe the finite element implementation of the constitutive SMA material model presented in Section 2.1, as well as the phase field coupled fracture/fatigue formulation described in Section 2.2, including details on damage irreversibility, strain energy decomposition and solution schemes. For simplicity, the temperature will be considered to be uniform throughout our numerical experiments but the extension to non-isothermal conditions is straightforward (see, e.g., [29,49,50]). The implementation is carried out in the commercial finite element package Abaqus by means of a user element (UEL) subroutine. Abaqus2Matlab is employed to pre-process the input files [51].

3.1. Implicit integration of the SMA constitutive model

The implementation of the constitutive model follows the work by Qidwai and Lagoudas [52], where the evolution of the transformation strain tensor is incrementally integrated using the backward Euler method. Thus, we use the common notation of adding the subscript $n + 1$ to quantities in the current time step while the subscript $n$ denotes the previous time step, e.g. for the transformation strain tensor:

\begin{equation}
\varepsilon_{n+1}^t = \varepsilon_n^t + \Delta \varepsilon_{n+1}^t.
\end{equation}

(26)
In addition, as the problem is solved in an iterative manner, the superscript \((k)\) is used for the iteration counter; i.e., (26) becomes
\[
e^{(k+1)}_{n+1} = e^{(k)}_{n+1} + \Delta e^{(k)}_{n+1}.
\]

The incremental stresses are computed using a return mapping algorithm; a purely elastic trial state is followed by a transformation correction stage. Thus, assuming a uniform temperature and denoting \(C\) as the elastic stiffness matrix, the elastic stress prediction is given by
\[
\sigma_{n+1} = \sigma_n + C \Delta \epsilon
\]

The return mapping algorithm is then used to enforce satisfying the transformation consistency condition \(\Phi = 0\). Specifically, the convex cutting plane algorithm [53] is used, which differs from the commonly used closest point projection algorithm as follows. Considering the flow rule (5) and (27), the incremental transformation tensor reads,
\[
\Delta \epsilon^{(k)}_{n+1} = \left( \xi^{(k+1)}_{n+1} - \xi_n \right) \Lambda^t \left( \sigma^{(k)}_{n+1} \right) - \left( \xi_{n+1}^{(k)} - \xi_n \right) \Lambda^t \left( \sigma^{(k)}_{n+1} \right).
\]

In the convex cutting plane algorithm the implicit dependence on the transformation direction \(\Lambda^t \left( \sigma^{(k)}_{n+1} \right)\) is relaxed, and (29) can be reformulated as,
\[
\Delta \epsilon^{(k)}_{n+1} = \Delta \xi^{(k)}_{n+1} \Lambda^t \left( \sigma^{(k)}_{n+1} \right).
\]

The total current strain is held constant during the iterative correction such that, considering (12), the incremental Cauchy stresses read,
\[
\Delta \sigma^{(k)}_{n+1} = -C^{(k)}_{n+1} \left( \Delta S \sigma^{(k)}_{n+1} + \Lambda^{(k)}_{n+1} \right) \Delta \xi^{(k)}_{n+1}.
\]

In an iterative setting, the consistency condition implies,
\[
\Phi^{(k)}_{n+1} + \Delta \Phi^{(k)}_{n+1} = \Phi^{(k+1)}_{n+1} \approx 0,
\]

such that applying of the chain rule and inserting (31) renders,
\[
\Phi^{(k)}_{n+1} - \partial \Phi^{(k)}_{n+1} \cdot C^{(k)}_{n+1} \left( \Delta S \sigma^{(k)}_{n+1} + \Lambda^{(k)}_{n+1} \right) \Delta \xi^{(k)}_{n+1} + \partial \Phi^{(k)}_{n+1} \Delta \xi^{(k)}_{n+1} \approx 0.
\]

And finally, solving (33) for the correction in the martensitic volume fraction at iteration \((k)\) gives,
\[
\Delta \xi^{(k)}_{n+1} = \frac{-\Phi^{(k)}_{n+1}}{\partial \Phi^{(k)}_{n+1} - \partial \Phi^{(k)} \cdot C^{(k)}_{n+1} \left( \Delta S \sigma^{(k)}_{n+1} + \Lambda^{(k)}_{n+1} \right)}.
\]

The transformation strain tensor can then be updated via (27) and (30). The updated values of the transformation strain and elastic stiffness are used to calculate an updated stress tensor via,
\[
\Delta \sigma^{(k)}_{n+1} = -C^{(k)}_{n+1} \left( \Delta S \sigma^{(k)}_{n+1} + \Lambda^{(k)}_{n+1} \right) \Delta \xi^{(k)}_{n+1},
\]

which is then used to compute the transformation function. The iterative procedure continues until \(\Phi^{(k+1)}_{n+1} \approx 0\) or \(\xi^{(k+1)}_{n+1}\) reaches the limiting values 0 or 1 [42].

The last step involves computing the consistent material Jacobian \(L\) [42,52,54], which for isothermal conditions is given by,
\[
d\sigma = L \, d\epsilon = \left( C + C \left( \Delta S \sigma + \Lambda^t \right) \otimes \frac{\partial \Phi}{\partial \Phi} \right) \, d\epsilon.
\]

### 3.2. Addressing irreversibility and crack growth in compression

Fracture is assumed to be driven by the total strain energy density \(\psi\); i.e., both the elastic \(\psi^e\) and transformation \(\psi^t\) strain energy densities contribute to cracking on equal footing. This choice is mainly phenomenological but has some physical background, as it appears sensible to assume a contribution from the transformation strains to the fracture process. Similar approaches have been adopted by other authors in relation to other inelastic quantities; see
e.g., Ref. [21,55] for examples in the context of plasticity and ductile damage. To maintain resistance in compression and during crack closure, the elastic contribution to the strain energy density is decomposed into volumetric and deviatoric parts, following Amor et al. [56]. In this approach, the deviatoric and tensile volumetric components contribute to fracture but the compressive volumetric term does not. Thus, the elastic strain energy is decomposed into the following two terms:

\[
\psi_+^e = \frac{1}{2} K (\xi) (tr (\epsilon))^2_+ + \mu (\xi) (\epsilon' : \epsilon') \\
\psi_-^e = \frac{1}{2} K (\xi) (tr (\epsilon))^2_-
\]

where \( K \) and \( \mu \) respectively denote the bulk and shear modulus, which would be dependent on \( \xi \) in the transformation region. This elastic strain energy decomposition is implemented following the hybrid approach by Ambati et al. [57]. One should note that the volumetric–deviatoric split is only applied to the elastic component of the strain energy density, such that

\[
\psi_+ = \psi_+^e + \psi'
\]

Secondly, a history variable field \( H \) is introduced to ensure damage irreversibility, \( \phi_{n+1} \geq \phi_n \). To ensure irreversible growth of the phase field variable, the history field must satisfy the Kuhn–Tucker conditions

\[
\psi_+ - H \leq 0, \quad H \geq 0, \quad H(\psi_+ - H) = 0
\]

for both loading and unloading scenarios. Thus, for a current time \( t \), the history field can be defined as

\[
H = \max_{t \in [0,1]} \psi_+ (\tau)
\]

3.3. Finite element discretisation

We proceed to formulate the two-field weak form of the problem and subsequently derive the stiffness matrices and residuals applying a finite element discretisation. Thus, consider the total potential energy of the solid, Eq. (19), under isothermal conditions and in the absence of body forces and external tractions. The first variation of (19) with respect to \( \epsilon \) and \( \phi \), gives

\[
\int_\Omega \left[ (1 - \phi)^2 \sigma : \delta \epsilon \right] dV = 0,
\]

\[
\int_\Omega \left[ -2(1 - \phi) \delta \phi H + G_c (\xi) \left( \frac{\phi}{\ell} \delta \phi + \ell \nabla \phi \cdot \nabla \delta \phi \right) \right] dV = 0.
\]

Now make use of Voigt notation and consider a 3D solid. The displacement field \( u \) and the phase field \( \phi \) are discretised as

\[
u = \sum_{i=1}^m N_i u_i \quad \text{and} \quad \phi = \sum_{i=1}^m N_i \phi_i
\]

where \( N_i \) is the shape function matrix, a diagonal matrix with \( N_i \) in the diagonal terms. Here, \( N_i \) denotes the shape function associated with node \( i \), \( m \) is the total number of nodes per element, and \( u_i = \{ u_x, u_y, u_z \}^T \) and \( \phi_i \) are the displacement and phase field values at node \( i \), respectively. Consequently, the corresponding derivatives can be discretised making use of the strain–displacement matrices \( B_i^u \) and \( B_i^\phi \) as follows:

\[
\epsilon = \sum_{i=1}^m B_i^u u_i \quad \text{and} \quad \nabla \phi = \sum_{i=1}^m B_i^\phi \phi_i.
\]

Here \( \epsilon = \{ \varepsilon_{xx}, \varepsilon_{yy}, \varepsilon_{zz}, \gamma_{xy}, \gamma_{xz}, \gamma_{yz} \}^T \), with \( \gamma \) being the engineering strain, such that \( \gamma_{xy} = 2\varepsilon_{xy} \).

We proceed to formulate the residuals and the stiffness matrices considering this finite element discretisation and the fact that (42)–(43) must hold for arbitrary values of \( \delta u \) and \( \delta \phi \). The associated discrete equations can be formulated as the following residuals with respect to the displacement field and the phase field variable, respectively,

\[
r_i^u = \int_\Omega \left\{ [(1 - \phi)^2 + \kappa] (B_i^u)^T \sigma \right\} dV
\]

\[
r_i^\phi = \int_\Omega \left\{ \phi (B_i^\phi)^T \sigma \right\} dV
\]
where \( \kappa \) is a numerical parameter introduced to keep the system of equations well-conditioned. A value of \( \kappa = 1 \times 10^{-7} \) is adopted throughout this work. Finally, the tangent stiffness matrices can be readily computed by taking the first derivative of the residual vectors, rendering

\[
K_{ij}^{uu} = \frac{\partial r_i^u}{\partial u_j} = \int_{\Omega} \left\{ \left[ (1 - \phi)^2 + \kappa \right] (B_i^u)^T \mathcal{L} B_j^u \right\} d\Omega
\]

(48)

\[
K_{ij}^{\phi\phi} = \frac{\partial r_i^\phi}{\partial \phi_j} = \int_{\Omega} \left\{ \left( 2H + \frac{G_c(\xi)}{\ell} \right) N_i N_j + G_c(\xi) \ell (B_i^\phi)^T B_j^\phi \right\} d\Omega
\]

(49)

### 3.4. Solution schemes

A global iterative scheme is adopted to obtain the displacement \( u \) and phase field \( \phi \) solutions for which \( r^u = 0 \) and \( r^\phi = 0 \):

\[
\left\{ \begin{array}{c}
u \\ \phi \end{array} \right\}_{t+\Delta t} = \left\{ \begin{array}{c}
u \\ \phi \end{array} \right\}_t - \left[ \begin{array}{cc} K^{uu} + M & 0 \\ 0 & K^{\phi\phi} \end{array} \right]^{-1} \left\{ \begin{array}{c} r^u \\ r^\phi \end{array} \right\}
\]

(50)

We develop a numerical implementation that can accommodate both staggered [43] and monolithic quasi-Newton schemes [23,58]. As shown by Wu et al. [58] for the PF-CZM model and by Kristensen and Martínez-Pañeda [23] for the AT2 model, the use of quasi-Newton methods such as the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm enables a robust implementation of monolithic schemes that retain unconditional stability, speeding up calculations by several orders of magnitude. In this work, this is particularly relevant for the analysis of fatigue, as otherwise calculations would be prohibitive [23]. The reader is referred to Refs. [43] and [23] for details on the implementation of staggered and monolithic quasi-Newton solution schemes, respectively.

### 4. Results

We proceed to demonstrate the potential of the computational modelling framework by addressing a number of case studies of particular interest. First, in Section 4.1, a boundary layer model is used to gain insight into the fracture behaviour of SMAs by investigating stationary and propagating cracks. Secondly, in Section 4.2, we proceed to model mode I fracture in a square plate, a paradigmatic phase field fracture benchmark. Mixed-mode conditions and crack coalescence are then investigated using an asymmetric double-notch bar in Section 4.3. Finally, we conduct a 3D large scale analysis of fatigue failure of a biomedical stent in Section 4.4.

Our numerical experiments are conducted on an equiatomic nitinol SMA, following the experimental data provided in Refs. [1,59]. The phase diagram transformation surface slopes for martensite (\( C_M \)) and austenite (\( C_A \)) are given at a reference stress of \( \sigma^* = 300 \) MPa. A uniform temperature of \( T = 320 \) K is generally adopted, following the experiments, but its influence will be investigated. In addition, we assume a uniform transformation strain, such that in (8), \( H = H_{\text{min}} = H_{\text{max}} \). The comparison between the calibrated model predictions and the experimental data from the uniaxial tension tests by Strnadel et al. [59] is shown in Fig. 2. The smooth hardening capabilities of the model enable attaining a very good agreement with the experiments.

Regarding the fracture behaviour, a value of 22.5 \( \text{kJ/m}^2 \) is adopted the material toughness from the range of reported data for NiTi [60], unless otherwise stated. However, one should note that this choice is based on the austenite fracture resistance, as a consequence of the assumption of a small scale transformation zone. The implications of this assumption will be discussed and results compared to the case of a martensite volume fraction-dependent critical energy release rate, \( G_r(\xi) \). In regard to the constitutive choices inherent to the phase field model, the conventional AT2 model is generally adopted, although calculations are also conducted with the AT1 model for comparative purposes. The material properties used are listed in Table 1.

#### 4.1. Boundary layer formulation

The concept of a boundary layer formulation is illustrated in Fig. 3. For a cracked solid, the crack tip stress state is characterised by the stress intensity factor; \( K_I \), assuming mode I conditions. The Williams [61] solution...
M. Simoes and E. Martínez-Pañeda


Fig. 2. Uniaxial tensile stress–strain response of NiTi showing the validation of the model against the experimental data by Strnadel et al. [59].

Table 1
Selected material parameters used in the numerical experiments, following the measurements by Strnadel et al. [59] on an equiatomic nitinol SMA.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Magnitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elastic properties</td>
<td></td>
</tr>
<tr>
<td>Austenite’s Young’s modulus, $E_A$ (MPa)</td>
<td>41 000</td>
</tr>
<tr>
<td>Martensite’s Young’s modulus, $E_M$ (MPa)</td>
<td>22 000</td>
</tr>
<tr>
<td>Austenite’s Poisson’s ratio, $\nu_A$</td>
<td>0.33</td>
</tr>
<tr>
<td>Martensite’s Poisson’s ratio, $\nu_M$</td>
<td>0.33</td>
</tr>
<tr>
<td>Phase diagram properties</td>
<td></td>
</tr>
<tr>
<td>Martensite start temperature, $M_s$ (K)</td>
<td>239</td>
</tr>
<tr>
<td>Martensite end temperature, $M_f$ (K)</td>
<td>221</td>
</tr>
<tr>
<td>Austenite start temperature, $A_s$ (K)</td>
<td>266</td>
</tr>
<tr>
<td>Austenite end temperature, $A_f$ (K)</td>
<td>282</td>
</tr>
<tr>
<td>$\sigma$ vs. $T$ slope (loading), $C_M</td>
<td>_{\sigma=300\text{MPa}}$ (MPa/K)</td>
</tr>
<tr>
<td>$\sigma$ vs. $T$ slope (unloading), $C_A</td>
<td>_{\sigma=300\text{MPa}}$ (MPa/K)</td>
</tr>
<tr>
<td>Other</td>
<td></td>
</tr>
<tr>
<td>Transformation strain $H$</td>
<td>0.0335</td>
</tr>
<tr>
<td>Material toughness $G_c$ (kJ/m$^2$)</td>
<td>22.5</td>
</tr>
<tr>
<td>Smooth hardening properties $n_1, n_2, n_3, n_4$</td>
<td>0.15, 0.17, 0.25, 0.15</td>
</tr>
<tr>
<td>Temperature $T$ (K)</td>
<td>320</td>
</tr>
</tbody>
</table>

for a linear elastic solid can be used to relate the displacement field to the magnitude of $K_I$. Considering a polar coordinate system $(r, \theta)$ and a Cartesian coordinate system $(x, y)$ centred at the crack tip, with the crack plane along the negative $x$-axis, the displacement solution reads:

$$u_i = \frac{K_I}{E_A} r^{1/2} f_i (\theta, \nu_A),$$

where the subscript index $i$ equals $x$ or $y$, and the functions $f_i (\theta, \nu)$ are given by

$$f_x = \frac{1 + \nu_A}{\sqrt{2\pi}} (3 - 4\nu_A - \cos \theta) \cos \left(\frac{\theta}{2}\right)$$

$$f_y = \frac{1 - \nu_A}{\sqrt{2\pi}} (3 - 4\nu_A - \cos \theta) \sin \left(\frac{\theta}{2}\right).$$
Fig. 3. Boundary layer concept, illustrated on a Compact Tension specimen.

and

\[ f_y = \frac{1 + \nu_A}{\sqrt{2\pi}} (3 - 4\nu_A - \cos \theta) \sin \left( \frac{\theta}{2} \right). \]  \hfill (53)

Thus, the nodal displacements in the outer boundary of the finite element model can be prescribed to evaluate
the crack tip behaviour at a given value of \( K_I \). However, one should note that this is under the assumption that the
inelastic region is small, generally referred to as small transformation zone conditions; analogous to the so-called
small scale yielding conditions in elastic–plastic materials. Accordingly, Eqs. (51)–(52) make use of the elastic
constants for the austenitic phase.

4.1.1. Stationary crack tip fields

The analysis of stationary cracks provides insight into the fracture behaviour of SMAs and facilitates interpre-
tation of the phase field fracture results. We adopt a boundary layer formulation, as described in Fig. 3, and take
advantage of symmetry to model only the upper half of the circular domain. After a mesh sensitivity study, a total of
14,183 quadrilateral quadratic elements with reduced integration are used. The mesh is refined close to the crack tip,
as shown in Fig. 4b. The loading is applied by prescribing a remote \( K_I \) field following Eqs. (51)–(52). Accordingly,
a reference length scale can be defined as [5]:

\[ L = \frac{1}{2\pi} \left[ \frac{K_I}{C_M(T - M_f)} \right]^2. \]  \hfill (54)

The finite element results obtained for a stationary crack in the reference NiTi material of Table 1 are shown
in Fig. 4. Consider first Fig. 4a, where the normalised crack tip stress distribution is shown as a function of the
normalised distance ahead of the crack tip. In agreement with expectations, a stress-induced transformation region
develops near the crack tip of SMAs. Three distinct domains can be observed. Adjacent to the crack tip, an inner \( K_I \)
field is observed within the martensitic region, where crack tip stresses exhibit the \( r^{-1/2} \) linear elastic singularity.
For distances larger than roughly 0.03\( L \) from the crack tip, a transformation region exists, where the stresses are
much less singular and the martensitic volume fraction is between 0 and 1, see Fig. 4b. Farther away from the crack
tip, the \( r^{-1/2} \) linear elastic singularity is again recovered, indicating the presence of an outer \( K_I \) field in the purely
austenitic region (\( \xi = 1 \)). For clarity, this outer \( K_I \) field associated with the austenitic phase is here frequently
denoted as \( K_\infty \). Depending on the material properties, the size of the sample and the temperature, the outer \( K_\infty \)
regime might be very small, which would complicate fracture mechanics testing—see Ref. [13] for a discussion.
Also, the consideration of plastic yielding will predictably introduce an additional crack tip region, adjacent to the
-crack tip and within the inner elastic domain. Fig. 4b shows the shape of the transformation zone, with blue and
red colours respectively denoting the martensitic and austenitic phases. It follows immediately from the existence
of this stress-induced transformation region that the \( J \)-integral becomes path-dependent; see, e.g. Ref. [62] for a
discussion on the path-dependence of \( J \) in inhomogeneous materials.
Fig. 4. Crack tip fields ahead of a stationary crack: (a) normalised crack tip stresses versus distance to the crack tip, and (b) contours of the martensite volume fraction $\zeta$. $L$ denotes the characteristic length scale associated with $K^\infty$, as given by Eq. (54).

4.1.2. Crack growth resistance curves (R-curves)

We proceed to model crack advance using the phase field fracture formulation described in Section 2.2. As in the stationary crack analysis, a boundary layer model is used. In this case, the refined region of the finite element mesh extends over the entire crack propagation domain, where in all calculations the characteristic element size is at least seven times smaller than the phase field length scale, following Ref. [29]. Approximately 47,200 quadrilateral linear elements are used. Crack growth resistance curves (R-curves) are predicted by computing the crack extension $\Delta a$ as a function of the remote load, as characterised by the remote stress intensity factor $K_f$. The remote load is normalised by a reference stress intensity factor, given by the following relationship:

$$K_0 = \sqrt{\frac{E_A G_c}{(1 - \nu_A^2)}}$$  \hspace{1cm} (55)
While the crack extension is normalised by a critical length $L_c$ [5], associated with the material toughness,

$$L_c = \frac{(1 - 2\nu_A)^2}{2\pi(1 - \nu_A^2)} \frac{E_A G_c}{[C_M(T - M_s)]^2}.$$  \hspace{1cm} (56)

The results, shown in Fig. 5, examine the influence of: (a) the phase field length scale, (b) the toughness definition, (c) the crack density function and (d) the temperature. In all cases, the model predicts a toughening effect (rising R-curve) associated with energy dissipation due to phase transformation, as observed in the experiments.

Consider first Fig. 5a, where the results reveal a rising R-curve with decreasing $\ell/L_c$ ratio. The trends observed can be explained as follows; recall (see, e.g., [29,44]) that the phase field length scale is related to the material tensile strength as,

$$\hat{\sigma} = \frac{9}{16} \sqrt{\frac{EG_c}{3\ell}}.$$  \hspace{1cm} (57)

Consequently, smaller values of $\ell$ lead to higher strengths and this results in greater inelastic dissipation, as it has been observed using cohesive zone models in SMAs [5] and in elastic–plastic materials [63,64]. However, note that in previous crack growth analyses in SMAs, the fracture energy was assumed to be constant. As shown in Fig. 5b, we proceed to define $G_c$ as a function of the martensite volume fraction using the rule of mixtures,

$$G_c(\xi) = (1 - \xi) G_c^A + \xi G_c^M,$$  \hspace{1cm} (58)
and evaluate its influence. Following Ref. [13], we take the toughness of the martensite phase \(G_c^M\) to be 20% smaller than that of the austenitic one, provided in Table 1. The results show that cracking initiates at a lower load level and this results in lesser dissipation. This is not surprising given the smaller magnitude of \(G_c^M\) and the fact that cracking initiates in the martensitic region. Note that \(K_0\) has been defined relative to the remote \(K_1\), using the elastic constants for the austenitic phase, see (55); this results in a higher prediction of the initiation load, which occurs when \(G = G_c\) is met locally. As shown in Ref. [65] for homogeneous elastic–plastic materials, the onset of crack growth in the presence of a large initial crack is based only on energy considerations; i.e., it occurs at \(K_1 = K_0\) and it is insensitive to the value of \(\ell\). In the SMA case, cracking initiates at \(K_I > K_0\) due to: (i) the differences between the inner and outer \(K\)-fields, as discussed in the stationary crack analysis, and (ii) the \(\xi\)-dependence of \(G_c\). The choice of a \(\xi\)-dependent fracture energy, not considered so far (see, e.g., [5,15,16]), appears to be a sensible one. Experiments show significant differences between the toughness of purely austenitic and purely martensitic samples [13], and the results obtained here (Fig. 5b) reveal non-negligible differences in the predicted crack growth resistance behaviour relative to the choice of a constant fracture energy.

Next, the influence of the constitutive choice for the phase field crack density function is assessed in Fig. 5c. The results reveal that cracking initiates earlier in the AT1 model and leads to less dissipation, relative to the AT2 case. Finally, the role of temperature is quantified in Fig. 5d. This is of interest because the reference temperature (320 K) is above the austenitic end temperature \(A_f\), which should lead to a full recovery in the wake of the crack. This is not the case for 253 K, which is below the austenitic start temperature \(A_s\), implying that no reverse transformation takes place upon unloading. The unloading response has proven to have an important effect in elastic–plastic materials, revealing big differences between isotropic and kinematic hardening laws [66,67]. In the case of SMAs, a higher degree of dissipation is observed in the cases with smaller temperature, where the magnitude of the austenite to martensite transformation stresses is smaller.

4.2. Fracture of a square plate with a crack

We proceed to model the failure of a cracked square plate subjected to tension, a paradigmatic benchmark in phase field fracture [23,29,43,57]. The specimen has an initial horizontal crack going from the left side to the centre of the specimen, both vertical and horizontal displacements are restricted in the bottom boundary, and we load the plate by prescribing the vertical displacement in the upper edge, \(u^\infty\). The geometric setup, dimensions (in mm) and boundary conditions are given in Fig. 6. The constitutive behaviour of the material is characterised by the parameters listed in Table 1 while the phase field model uses \(G_c = 4.1 \text{ kJ/m}^2\) and \(\ell = 0.0075\ mm\). The characteristic element size is at least 7 times smaller than \(\ell\) along the extended crack plane. A total of 45,571 quadrilateral linear elements with full integration are used.

The results obtained are shown in Fig. 7 in terms of the force versus displacement response, with contours of martensite volume fraction \(\xi\) and phase field fracture parameter \(\phi\) at different stages embedded in the figure. Differences from the response commonly observed in this classic benchmark are notable. In a linear elastic homogeneous material the plate fails in an unstable manner, with the force versus displacement curve exhibiting a very large drop immediately after reaching the peak load—see, e.g. [23,43]. Contrarily, in the SMA sample a significant toughening effect is observed; there is an initial drop in the load associated with the first instance of crack growth but then the crack progresses in a stable manner until complete failure of the sample. The toughening effect observed is undoubtedly related to the energy dissipated due to transformation. As shown in Fig. 7, a large transformation zone develops in the sample, exhibiting a shear banding-like behaviour that resembles that observed in elastic–plastic materials [68]. However, one should note that, for the material properties here considered, the crack still follows the mode I fracture path.

4.3. Plane strain tension of an asymmetric double-notched specimen

An asymmetrically notched plane strain bar is investigated to model mixed-mode fracture and crack coalescence. The double-notched bar, depicted in Fig. 8, is clamped at the bottom end \((u_x = u_y = 0)\) and subjected to a vertical displacement \(u^\infty\) at the top edge. Two circular notches of radii 2.5 mm have been geometrically introduced. The bar is assumed to be made of the equiatomic nitinol SMA whose properties and model parameters are listed in Table 1. The phase field length scale equals 0.2 mm and the finite element mesh is chosen accordingly, with the
characteristic element length in the region between the two notches being on the order of 0.05 mm. A total of 23,622 quadrilateral plane strain linear elements are used to discretise the model.

The finite element results computed are shown in Fig. 9, including the force versus displacement response as well as contours of phase transformation and phase field fracture parameter. Contrarily to the response observed in the previous case study, Section 4.2, a sharp drop in the force versus displacement curve is observed, indicative of brittle fracture with little inelastic dissipation. This can be rationalised by observing the phase transformation...
Fig. 8. Plane strain tension of an asymmetric double-notched specimen: dimensions (in mm) and loading configuration.

Fig. 9. Plane strain tension of an asymmetric double-notched specimen: force versus displacement response. The figure includes contours of martensite volume fraction $\xi$, just before failure, and phase field fracture parameter $\phi$, after the unstable crack growth event.

contour just before failure, at $u^\infty = 0.39$ mm; as shown in Fig. 9, the inelastic region is confined to a small area in the close vicinity of the tips of the notches. No inelastic shear bands are observed such that as soon as cracking initiates at the notch tips, the cracks coalesce and unstable cracking is observed. We note that this finding is specific to the boundary value considered, a parametric analysis is needed to characterise the interplay between the different scales at play (notch radius, phase field length scale, sample dimensions) and its implications on fracture stability due to inelastic dissipation.

4.4. Fatigue failure of a NiTi stent

The capabilities of the modelling framework in capturing fatigue crack growth are demonstrated by simulating cyclic damage in a stent, a paradigmatic application of shape memory materials [69–71]. This case study also serves
to showcase the computational efficiency of the framework and its applicability to the modelling of computationally demanding large-scale 3D boundary value problems.

Stents are small cylindrical tubes that are placed into blood vessels, arteries, or other ducts to hold the structure open. Often, their role is to counteract the effects of vascular diseases that are associated with plaque blockages that hinder fluid flow, see Fig. 10. To deploy the stent, it is usually crimped, placed in a delivery system (e.g., catheter), and finally expanded in vivo to widen the duct. Nitinol is a popular material choice in stent manufacturing due to its biocompatibility and capacity to expand by recovering its elastic deformation after the constraining delivery system has been removed (superelasticity). However, fatigue resistance is often the limiting design criterion as the stent is subjected to repeated contraction and expansion during the systolic and diastolic cycles. Current fatigue design models for stents commonly use Goodman’s and other empirical methods to estimate the number of cycles to failure by extrapolating from the stress/strain state of the first cycle. We aim here at providing a more mechanistic approach using a phase field fatigue framework that can predict features such as S–N curves or the Paris law as a natural outcome of the model [35].

We assume that the stent has been manufactured using the equiatomic NiTi whose material properties are listed in Table 1. Taking advantage of symmetry along the longitudinal direction, half of the stent is modelled. The stent is subjected to different expansion and compressive pressures by prescribing a radial displacement in the outer surface, \( u_r \). First, there is a crimp phase where the simulation emulates the compression of the stent inside of the capsule prior to delivery; a radial displacement of \( u_r = -0.16 \) mm is applied. Secondly, the stage of deployment is reproduced by allowing the stent to expand up to \( u_r = -0.02 \) mm, where further expansion is limited by the surrounding duct. Once deployed, the stent is subjected to cyclic loads of amplitude \( \Delta u_r = -0.04 \) mm that simulate the compression and expansion pressures experienced due to the systolic and diastolic cycles. The damage response will be governed by the phase field fatigue model described in Section 2.2.1, with a phase field length scale of \( \ell = 0.02 \) mm and using the monolithic quasi-Newton solution scheme. The magnitude of \( \ell \) is at least 2.5 times larger than the characteristic element length. Eight-node brick elements with full integration are used to discretise the geometry, with the model containing more than 7 million degrees-of-freedom (DOFs). Calculations are run in parallel, using 16 cores, with each load increment taking approximately 2 min.

The finite element results obtained are shown in Figs. 11–13, along with a video that is provided in the online version of this manuscript (Video 1). Fig. 11 shows the deformed shape of the model during the four stages of the analysis: the crimp and expansion stages of the deployment phase, and the systolic compression and diastolic expansion stages associated with each pressure cycle. Contours of the phase field are shown with the deformed shape, revealing that values of \( \phi \) of up to 0.56 are attained during the deployment phase. Thus, a significant amount of damage occurs during deployment but no cracking is observed. This is also the case for the phase transformation; most of it takes place during the deployment phase. The contours of martensitic volume fraction \( \xi \) at the end of the crimp–expansion deployment process are shown in Fig. 12.
Fig. 11. Phase field contours and deformed shape (×8) of the SMA stent during crimp, deployment and the first systolic–diastolic pressure cycle.

The ξ contours show that the martensite phase is localised in the edges of the stent struts. This is also the location where cracking initiates. While φ reaches high values during the deployment stage, 29 cycles of (compression–expansion) systolic–diastolic pressure are needed for cracks to initiate. As shown in Fig. 13, these surface cracks initiate in the regions where phase transformation took place during stent deployment. The cracked region (φ = 1) extends with increasing the number of pressure cycles, and after 50 cycles it has extended over a significant part of the stent, including the bridging areas between struts. The evolution of the phase field over 6 cycles is shown in Video 1, provided in the online version of this manuscript. In summary, the example demonstrates that the framework presented here can be used to estimate the lifetime of medical stents for arbitrary geometries, boundary conditions and material properties.

Fig. 12. Contours of the martensitic volume fraction ξ in the SMA stent during the deployment phase.
5. Conclusions

We have presented the first phase field fracture formulation for Shape Memory Alloys (SMAs). The model can capture both fracture and fatigue damage and includes the main features inherent to the constitutive behaviour of SMAs; including superelasticity, shape memory effect, gradual phase transformations and the stress-sensitivity of the inelastic recoverable strain and of the phase diagram slope. From a fracture perspective, key features include the definition of a martensite volume fraction ($\xi$)-dependent fracture energy and the consideration of the AT1 and AT2 models in the constitutive definition of the crack density function. The theoretical model is numerically implemented using the finite element method, including both staggered and (quasi-Newton) monolithic schemes. The potential and robustness of the computational framework presented are demonstrated by addressing several paradigmatic 2D and 3D boundary value problems involving subcritical crack growth, unstable cracking, crack coalescence and fatigue crack growth of a nitinol stent. The main findings are:

(i) A stress-induced transformation zone develops in the vicinity of the crack tip and three distinct regions are observed: an inner martensite region with $r^{-1/2}$ singularity, an intermediate phase transformation region, and an outer austenite region with $r^{-1/2}$ singularity.

(ii) The stress-induced transformation phenomenon leads to inelastic energy dissipation and material toughening. This toughening effect is more significant at higher material strengths and lower temperatures. The use of a uniform toughness and the AT2 choice for the crack density function also increases crack growth resistance relative to a $\xi$-dependent fracture energy and the AT1 model, respectively.

(iii) Boundary value problems that favour the appearance of inelastic shear bands and large phase transformation regions can lead to notable subcritical crack propagation. Contrarily, where the phase transformation is confined to a small region fracture occurs in an unstable manner.

(iv) The coupling with very recent developments in phase field fatigue enables predicting the service life of SMA components in practical applications, as demonstrated with the analysis of a stent.

Potential extensions to the current framework include the consideration of plastic straining and of thermal loads.

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.
Acknowledgements

The authors would like to acknowledge the assistance of Dr Darren J. Hartl (Texas A&M University) in relation to the implementation of the SMA constitutive model. M. Simoes acknowledges insightful discussions with Dr Advenit Makaya (European Space Agency, ESA) and Dr Christopher Braithwaite (University of Cambridge), as well as financial support from the EPSRC, UK (grant EP/R512461/1) and from the ESA, France (Contract No. 4000125861). E. Martínez-Pañeda acknowledges financial support from the EPSRC, UK (grants EP/R010161/1 and EP/R017727/1) and from the Royal Commission for the 1851 Exhibition, UK (RF496/2018).

References

M. Simoes and E. Martínez-Pañeda


