Explicit finite element implementation of an improved three dimensional constitutive model for shape memory alloys

A.P. Stebner a, L.C. Brinson a,b

a Department of Mechanical Engineering, Northwestern University, Evanston, IL 60208, USA
b Department of Materials Science and Engineering, Northwestern University, Evanston, IL 60208, USA

A R T I C L E  I N F O

Article history:
Received 19 September 2012
Received in revised form 20 November 2012
Accepted 27 December 2012
Available online 16 January 2013

Keywords:
Shape memory alloy
Constitutive model
Numerical implementation
Phase transformation

A B S T R A C T

This article documents a new implementation of a three dimensional constitutive model that describes evolution of elastic and transformation strains during thermo–mechanical shape memory alloy loading events assuming a symmetric, isotropic material response. In achieving this implementation, improvements were made to the original formulation of the constitutive model. These improvements allow for robust three-dimensional calculations over a greater range of thermo-mechanical loadings. Furthermore, a new explicit scheme for solving the model equations was derived. This scheme removed the need for user calibration of the numerical integration parameters and greatly reduced the sensitivity of this explicit finite element implementation of a rate independent model to mass scaling. Studies were performed that quantified both simulation times and convergence of the new scheme along with the original solution scheme of Panico and Brinson for single element and multi-element simulations. The effectiveness of the new scheme is apparent in 6 and 30 times reductions in computation expense for selected single and multi element simulations, respectively.

1. Introduction

1.1. Motivation

The Panico–Brinson constitutive model [1] has become a standard for comparison in modern day development of three-dimensional continuum-scale shape memory alloy (SMA) constitutive models (e.g., [2–5]), largely because it was the first model at this scale to phenomenologically consider the ability of a martensite microstructure to reorient during non-proportional loading events. In the years since original publication, several numerical implementations for modeling the SMA behaviors described by the thermodynamic formulation of Panico and Brinson [1] have been developed [1–3,6,7]. However, the original implementation scheme [1,6] is cumbersome to code and limited in numerical robustness. In this article, we proceed to present a new numerical implementation, which is comprised of:

- An improved constitutive model that allows for robust performance under a wider range of thermo-mechanical loadings while simplifying the set of model inputs;
- A new explicit integration scheme that is easier to implement, easier to use, and greatly reduces the computational expense when compared with the original Panico and Brinson implementation of this model.

This new numerical implementation serves to document robust convergence of the SMA constitutive law within structural simulations using explicit finite elements, and also improve the ability of other researchers to continue to benchmark new advancements in SMA constitutive modeling. Furthermore, the self-calibrating numerical scheme we derive provides a simple, physical, robust solution to solving constitutive equations that describe mechanics of reversible inelastic deformation governed by a bounded internal variable. The numerical complications inherent to such a constitutive model have been a recent topic of discussion (e.g. [4]).

Expanding upon the first point listed above, while the chosen form of the orientation kinetic function (see Section 1.2 for terminology definitions) of the Panico–Brinson model was theoretically robust [1], it was not sufficient under large mechanical loads when limited by machine precision (see Section 2.2.1). Thus, a new function has been introduced to describe the evolution of the oriented martensite volume fraction. Additionally, the number of model inputs has been reduced while retaining the original material physics captured by the Panico–Brinson formulation. Finally, Arghavani et al. [2] formulated a constitutive model with decoupled sets of equations describing transformation strain: one for orientation, a
second for reorientation. Inspired by the computational impact of [2], in the newly formulated numerical implementation we arrive at two smaller sets of equations describing the transformation strain that may be independently solved, even though they are not fully decoupled in formulation as in [2], and this simplifies the numerical implementation.

To the second point, Panico and Brinson originally proposed an implicit integration scheme (Newton–Raphson) for solving the partial differential equations (PDEs) of their model at integration points of explicit finite elements (FEs) [1]. However, this scheme requires programming the Jacobian of the equations comprising the model, which is tedious and error-prone due to the large number of equations to derive and code. This originally proposed scheme and the Jacobian for the improved constitutive model presented in this article are given in Appendix B for consideration in this light. In creating the first Abaqus Explicit [8] user material subroutine (VUMAT) of the proposed numerical implementation of the Panico and Brinson model, the Newton–Raphson integration was augmented with an explicit constant Lagrange multiplier integration (“Scheme 1” in this article, see Section 3.1) that overrode the implicit calculation in the event Newton–Raphson failed to converge within a few iterations. The original explicit override VUMAT was used in the Abaqus calculations published in [6,7].

In this article, complexity, convergence, and computational efficiency of two VUMATS that make calculations using unique explicit integration schemes will be compared: (1) the constant Lagrange multiplier integration scheme used to override the implicit integration scheme in the original Panico and Brinson VUMAT (Scheme 1, see Section 3.1) and (2) a newly derived explicit integration scheme that allows for more robust computations (Scheme 2, see Section 3.2), which has been implemented into a new VUMAT—the second written by our research group. Comparisons will be made within the context of non-proportional loading of a single three-dimensional brick element, as well as loading of a SMA bar containing a hole, modeled with a two-dimensional mesh of non-uniform elements.

We proceed with a brief review of the background necessary to understand the model features. Subsequently, the improved constitutive model is presented in Section 2, and in Section 3 the numerical schemes used in the two VUMATS are presented, including the new explicit integration scheme formulated for integrating the coupled equations of the improved model. Both single and multi element simulation data are then used to compare and contrast the convergence and computational efficiency of the VUMATS in Section 4. Finally, implications of these results are discussed in the conclusion.

1.2. Background

The shape memory behaviors exhibited by SMAs result from a reversible solid-state phase transformation between an austenite phase comprised of a high symmetry crystal structure and a lower symmetry martensite phase. A simple cartoon of crystal structure and microstructure changes during phase transformation is shown in Fig. 1. As indicated in Fig. 1, in this work forward describes austenite to martensite phase transformation while reverse indicates martensite to austenite. First let us consider differences in crystal structure between the phases (Fig. 1a–c). In this example the austenite square (Fig. 1a) transforms to martensite parallelograms (Fig. 1b and c). Note it is possible to impart two unique deformations upon the high symmetry square that result in the same low symmetry parallelogram when viewed from the perspective of each local coordinate system defined by the axes of Fig. 1b and c. Note that if we restrict the example of Fig. 1 to two dimensions in the plane of the paper, the two martensite parallelograms are not frame indifferent [9,10]; that is Fig. 1b cannot be rotated in the plane of the page to be of the same orientation as Fig. 1c. In this work, each unique stretch that may result from transformation of a single austenite crystal orientation defines a unique variant of martensite. Indeed, in a real material in three-dimensional space there are more than two unique variants of martensite – in the most prevalent SMA, NiTi, there are twelve monoclinic variants that may transform to and from a single body centered cubic austenite orientation [11,12].

Next let us consider microstructures that form from transformation events (Fig. 1d–f). The amount of material that transforms from austenite to a given martensite variant is driven by energy minimization [12]. In the absence of an applied load, mixtures of variants form in twinned arrangements that self-accommodate to preserve the macroscopic shape, since macroscopically there is no driving force for deformation. This is depicted in transforming from the microstructure comprised of repeating units of square austenite crystals in Fig. 1d to the martensite microstructure that is a mixture of two parallelogram variants in Fig. 1e. Contrarily, if the material transforms in the presence of an applied mechanical load, forming a martensite microstructure of variants whose orientations best accommodate the applied load will minimize the energy of the system. In the simple cartoon, transforming from the austenite microstructure (Fig. 1d) to a martensite microstructure of a single parallelogram variant under a tensile load best accommodates the macroscopic driving force for overall elongation (Fig. 1f). Here we have provided a high-level, over-simplified summary of the concepts of SMA microstructures during phase transformation to motivate the terminology and features of our model. The reader is referred to [12,13] and references therein for further reading on self-accommodation as well as other microstructural terms used in this work.

Fig. 1. A simplified, two-dimensional cartoon of crystal structures and microstructures associated with the reversible, solid-state phase transformation that gives rise to the shape memory behaviors of SMAs. Their significance is discussed in the text.
responses associated with shape memory behaviors are shown in Fig. 2 along with our simplified cartoons of microstructure at the inflection points of the macroscopic responses. Stress-induced phase transformation (Fig. 2a) may be exhibited by SMAs that are austenitic at ambient temperature and subjected to changes in mechanical load. Application of the load drives forward transformation, release of the load stimulates reverse phase transformation, and ideally recovery of the inelastic strain imparted upon loading is complete, thus stress-induced shape memory behavior is often called pseudo-elasticity or superelasticity. The plateaus in the stress–strain curve signify inelastic deformation resulting from the microstructure changing from the austenite to an oriented martensite and back again.

Thermally-induced phase transformation is achieved by heating and cooling the alloy above and below its characteristic transformation temperatures. Four temperatures typically describe the phase transformation: (1) martensite start and (2) martensite finish temperatures signify the onset and saturation of the forward transformation upon cooling; analogously (3) the austenite start and austenite finish temperatures characterize the reverse transformation upon heating. Further described in the caption of Fig. 2 and depicted in Fig. 2b, cooling from austenite to self-accommodated martensite load-free (microstructurally Fig. 1d to e), loading martensite from a self-accommodated state to an oriented state (Fig. 1e–f), unloading the oriented martensite (stays Fig. 1f), then heating back to austenite (Fig. 1d–f) is called the shape memory effect (SME). Thermally cycling under load between austenite and oriented martensite (Fig. 1d–f and back to d) gives rise to actuation.

Having briefly reviewed shape memory behaviors at a level required to understand the model, this introduction concludes with a brief review of the Panico–Brinson model followed by a definition of the terminology used to describe the model features, which has been modified from that used in the presentation of the original model[1]. Note that shorthand, operators, and abbreviations as well as their correlation to that used in[1] are defined in Appendix A.

The constitutive assumptions made in deriving the Panico–Brinson model are [1]:

(1) The total strain may be additively decomposed into elastic (e) and transformation (tr) strains:

\[ \varepsilon = \varepsilon_e + \varepsilon_{tr}. \] (1)

Furthermore, the rate of transformation strain may be decomposed into orientation (\( \sigma \)) and reorientation (re) parts:

\[ \dot{\varepsilon}_{tr} = \dot{\varepsilon}_e + \dot{\varepsilon}_{tr} \] (2)

(2) The transformation strain is traceless since there is less than 1% volume change exhibited by most SMAs in transforming from austenite to martensite crystal structures ([12]):

\[ \text{tr}(\varepsilon_{tr}) = 0. \] (3)

(3) The martensite volume fraction \( \xi \) may be expressed as a sum of self-accommodated (SA) and oriented (\( \sigma \)) martensite:

\[ \xi = \xi_{SA} + \xi_{\sigma}. \] (4)

and martensite volume fractions are bounded by zero and unity:

\[ 0 \leq \xi_{SA}, \xi_{\sigma} \leq 1. \] (5)

(4) Making the simplifying assumption that the material is isotropic and symmetric, the transformation strain is bounded by the magnitude of maximum polycrystalline transformation strain \( \varepsilon_{max} \).

(5) Furthermore, the magnitude of transformation strain is assumed to only evolve through orientation of martensite, independent of reorientation. Considering this assumption along with assumption (4), the evolution equation for \( \xi_{\sigma} \) is:

\[ \dot{\xi}_{\sigma} = \frac{\varepsilon_{tr} \cdot \dot{\varepsilon}_{tr}}{\sqrt{3/2} \varepsilon_{max} \| \dot{\varepsilon}_{tr} \|}. \] (6)

Another consequence of these assumptions is that reorientation may only re-direct the transformation strain, not change its magnitude, thus the tangency condition must always be satisfied between the transformation strain and the reorientation strain rate:

\[ \varepsilon_{\sigma} : \varepsilon_{re} = 0 \] (7)

(6) Rate and history independent kinetics are also assumed.

The model derivation is achieved by writing an expression for the Helmholtz free energy incorporating these constitutive assumptions:

\[ \psi(\varepsilon_e, T, \xi_{\sigma}, \xi_{SA}) = \frac{1}{2\rho} \varepsilon_e : C : \varepsilon_e + \xi_{\sigma}(T\theta_0 - u_0) + \xi_{SA}(T\theta_0 - u_0) + \xi_{\sigma}C_{\xi_{\sigma}} \left\{ (T - T_0) - T\ln \left( \frac{T}{T_0} \right) \right\} + \frac{1}{2} H_e \xi_{\sigma}^2. \] (8)

Fig. 2. Shape memory behaviors: (a) Pseudo-elasticity is exemplified by large inelastic strains induced during loading that are recovered during unloading due to reversible stress-induced phase transformation between austenite and martensite. (b) In the shape memory effect (SME), inelastic strains are induced during loading due to orientation of martensite variants from self-accommodated microstructure (black curve); residual strain upon unloading is still recoverable by phase transformation of this oriented martensite into austenite via heating (red curve, bottom path). Subsequent stress-free cooling from austenite to martensite (blue curve, Temperature axis) results in a self-accommodated microstructure – this process exhibits no macroscopically observable deformation. Actuation may be achieved via a special case of the SME in which the applied load is not removed from the alloy as it is thermally cycled through phase transformation between oriented martensite and austenite microstructures (red and blue curves, top path). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
This free energy is then examined in the Clausius–Duhem inequality to arrive at the following dissipation potential, where \( \mathbf{X} \) denotes a driving force:

\[
\dot{\rho}_{D_p} = \mathbf{X}_\sigma : \dot{\varepsilon}_\sigma + \mathbf{X}_{\varepsilon_r} : \dot{\varepsilon}_{\varepsilon_r} + \mathbf{X}_{\Delta} \dot{\varepsilon}_{\Delta} \geq 0.
\]

The driving forces are:

\[
\mathbf{X}_\sigma = \sigma' - \frac{\rho(T_0 - \bar{u}_0 + H(\varepsilon_{\varepsilon_r}))}{\sqrt{3/2}C_{\text{max}}} \mathbf{N}(\varepsilon_{\varepsilon_r})
\]

\[
\mathbf{X}_{\varepsilon_r} = \sigma' ;
\]

\[
\mathbf{X}_{\Delta} = -\rho(T_0 - \bar{u}_0).
\]

(9)

Evolution equations for the internal variables:

\[
\dot{\varepsilon}_\sigma = \dot{\varepsilon}_\sigma, \quad \dot{\varepsilon}_{\varepsilon_r} = \dot{\varepsilon}_{\varepsilon_r} ; \quad \mathbf{X}_\sigma ;
\]

\[
\dot{\varepsilon}_{\Delta} = \dot{\varepsilon}_{\Delta} ; \quad \mathbf{X}_{\Delta}.
\]

(11)

where \( \mathbf{N}(\varepsilon_{\varepsilon_r}) = \varepsilon' - N(\varepsilon_{\varepsilon_r}) \mathbf{N}(\varepsilon_{\varepsilon_r}) \), and the kinetic relations are introduced through limit functions (F):

\[
F_\sigma = |\mathbf{X}_\sigma| - Y_\sigma \dot{\varepsilon}_\sigma
\]

\[
F_{\varepsilon_r} = |\mathbf{X}_{\varepsilon_r}| - Y_{\varepsilon_r} \dot{\varepsilon}_{\varepsilon_r}
\]

\[
F_{\Delta} = \mathbf{X}_{\Delta} - Y_{\Delta} \dot{\varepsilon}_{\Delta}
\]

which are subject to Kuhn–Tucker conditions:

\[
F_\sigma \leq 0; \quad \Delta_\sigma \geq 0; \quad \Delta_\varepsilon_{\varepsilon_r} F_{\varepsilon_r} = 0
\]

(12)

where \( \sigma = \sigma, \varepsilon_r, \Delta, \mathbf{X}_\sigma \).

The functions \( Y_\sigma \) are the kinetic equations; these are documented in the presentation of the implementation of the improved model in Section 2.1.

The updated terminology used in this presentation provides a complete description of the macroscopic thermodynamic responses that are modeled while more accurately correlating with the microstructural phenomena that give rise to the shape memory behaviors.

- **Self-accommodated martensite** ("twinned" in [1]) transforms to and from austenite in a macroscopically shape-preserving manner, corresponding to transformation between austenite and self-accommodated martensite microstructures (Fig. 1d to and from Fig. 1e).
- **Oriented martensite** ("detwinned" in [1]) forms from self-accommodated martensite via an applied load (Fig. 1e and f), and also during phase transformation under load (Fig. 1d to and from Fig. 1f). Formation of oriented martensite gives rise to transformation strain, and in transforming oriented martensite back to austenite the transformation strain is recovered.
- **Orientation** ("transformation" in [1]) encompasses growth or depletion of oriented martensite. This model feature always controls the magnitude of transformation strain. In a case of proportional loading, this feature also controls the transformation strain direction. Note that "orientation" in this sense may mechanistically refer to:
  1. Phase transformation under load (austenite to/from oriented martensite, Fig. 1d to and from Fig. 1f);
  2. Conversion of self-accommodated martensite to oriented martensite (Fig. 1e to and from Fig. 1f);
  3. Furthermore, "orientation" does not include austenite-self-accommodated martensite phase transformation (Fig. 1d to and from Fig. 1e) as this process does not affect the macroscopic transformation strain.
- **Reorientation** encompasses changes in orientation, but not the magnitude of transformation strain during a non-proportional mechanical loading event. This model mechanism reflects the ability of an SMA to reorient its martensite variants to form a newly oriented microstructure as proportions of loading conditions change. This feature was the primary novelty of the Panico–Brinson model at the time of publication, though models published around the same time and since also include this ability (e.g., [2–4,14,15]).

Relative to the original publication, this terminology is more precise since orientation of self-accommodated martensite (Fig. 1e and f) is not a phase transformation process, while transformation between austenite and self-accommodated martensite (Fig. 1d to and from Fig. 1e) is a phase transformation process. The total inelastic strain modeled by both the original and improved formulations is transformation strain—the "reorientation" strain described below is recoverable through phase transformation. Thus it provides better agreement to call the total inelastic strain "transformation strain" and to allow it to still evolve by two mechanisms—"orientation" and "reorientation."

Hence, both the improved model presented in this work and the original Panico–Brinson model simulate three-dimensional evolution of elastic and transformation strains during pseudo-elastic, shape memory effect, and actuation events. They also provide the ability for the transformation strain to reorient in a tangential manner relative to the transformation surface during non-proportional loading events (Eqs. (7) and (10)). Furthermore, the models are able to simulate independent energy dissipation events for each formation of oriented martensite as well as reorientation of oriented martensite (Eqs. (9) and (12)). The current form of the model still assumes symmetric and isotropic material responses and neglects plasticity, thermal expansion, finite deformation, minor loop kinetics, cyclic evolution, and rate dependence. All features that may be added to this core model in future development efforts. The reader is referred to [3–5,14–29] and references therein for examples of advancements in phenomenological modeling of these mechanisms. Subsequent to original publication, Panico incorporated a mechanism to model irreversible strains that result from retained martensite [6,7], a mechanism originally proposed by Malecot et al. [27]. This unrecovered strain mechanism is not reviewed here nor included in the improved model since the present focus is on numerical implementation of the core constitutive model (elasticity + transformation), not the mechanics of unrecovered inelastic strains during cycling.

Background in hand, we proceed with documenting the improved model and the new numerical implementation.

2. The improved model

The model introduced in this section is improved from the original formulation [1]. As mentioned in the introduction, it has been found that equations, primarily the martensite volume fraction kinetic equations, must be modified to achieve robust three-dimensional computations. In making these modifications, the number of model inputs may also be reduced from 17 to 14. Additionally, since it was assumed that reorientation does not change the magnitude of the internal variables (inelastic strains and martensite volume fractions), but merely reorients the inelastic strain direction with respect to the transformation strain limit surface [1], it is possible to numerically solve for reorientation and orientation independently [2]. The fundamental mechanics and constitutive assumptions of the original formulation are not altered in this new formulation.

2.1. The model & solution algorithm

Here, the improved model is presented through the rate independent algorithm used to solve the equations within the VUMATs. We then document the improvements in detail in Section 2.2.
1. Abaqus prescribes increments \( \Delta e \) and \( \Delta T \) to the VUMAT, along with values of variables from the previous time step (denoted with superscript \( n \)). Strain and temperature are updated:

\[
T = T^n + \Delta T
\]

\[
e = e^n + \Delta e
\]

(14) (15)

2. The deformation resulting from these increments is assumed to be purely elastic. A trial stress is calculated using:

\[
\sigma_{\text{trial}} = C : (e - e^n)
\]

(16)

3. A check is made to see if this assumption of elastic deformation violates the Kuhn–Tucker conditions imposed on the limit functions of orientation and reorientation mechanisms. This check is achieved by first calculating the thermodynamic driving force for orientation:

\[
X_\sigma = \sigma_{\text{trial}} - \frac{\rho(\bar{T}^0 - \bar{u}_0) + H_{\text{tr}} e^n}{\sqrt{3/2}e^n_{\text{max}}} N(e^n)
\]

(17)

Next, the limit functions are evaluated:

\[
F_{\sigma} = ||X_\sigma|| - Y_{\sigma}(e^n)
\]

(18)

\[
F_{re} = ||\tilde{\tau}(e^n) : \sigma_{\text{trial}}|| - Y_{re}
\]

(19)

The form of the projection tensor \( \tilde{\tau} \) originally used by Panico and Brinson is \( \tilde{\tau}(\varepsilon) = \tau - N(\varepsilon)N(\varepsilon) \). The deviatoric identity tensor \( I \) is defined in full matrix indicial notation as \( I_{ijkl} = \delta_{ij} \delta_{kl} - \frac{1}{3} \delta_{ik} \delta_{jl} \) where \( i, j, k, l \) are 6 \times 6 identity tensor and \( \delta \) is Kronecker’s delta, or in Voigt notation as \( I^{(e)} = 1 - \frac{1}{2} I \) where \( I \) is the traditional 2nd order identity tensor expressed as a 1 \times 6 column vector. While the two forms of the projection tensor are not equivalent, in implementing this model one may also choose to replace the deviatoric identity simply with a fourth order identity tensor such that \( \tilde{\tau}(e^n) = I - N(e^n)N(e^n) \) since in the model \( \tilde{\tau}(e^n) \) is always contracted with \( \sigma'e \) and \( \tau - N(e^n)N(e^n) : \sigma' \). This result follows from a standard identity of deviatoric tensors, as described on page 423 of [30].

The kinetic scalar \( Y_{\sigma} \) represents an internal energy (friction) that must be overcome for the SMA microstructure to reorient, and is assumed to be constant in its phenomenological average. It may be calibrated via fitting to non-proportional loading data for a given material. This kinetic dissipation parameter provides hysteretic in reorientation, as a real material would experience. It is assumed to be constant merely because the phenomenological dissipation of a conglomerate of reorienting martensite variants is not quantified well enough to understand the true physical form. However, a physically motivated scalar function may easily be introduced in place of this parameter once this phenomenon is understood.

Analogously, \( Y_{re} \) controls the kinetics of \( \tilde{\sigma} \) and in this improved model takes the form:

\[
Y_{re}(\tilde{\sigma}) = \begin{cases} 
\tilde{\sigma} + \tan \left( \pi (\tilde{\sigma} + \frac{1}{2}) \right) + C' & \Delta \tilde{\sigma} > 0 \\
(1 - \tilde{\sigma}) (A' + \tan \left( \pi (\frac{1}{2} - \tilde{\sigma}) \right)) + C' & \Delta \tilde{\sigma} < 0
\end{cases}
\]

(20)

and the direction of \( \Delta \tilde{\sigma} \) is evaluated through:

\[
X_{\sigma} : N(e^n) > 0, \quad \Delta \tilde{\sigma} > 0
\]

\[
X_{\sigma} : N(e^n) < 0, \quad \Delta \tilde{\sigma} < 0
\]

(21)

If the conditions

\[
F_{\sigma-re} \leq 0
\]

(22)

are satisfied, then the assumption of the deformation prescribed by Abaqus being completely elastic is valid. \( \sigma = \sigma_{\text{trial}} = e^n \), and we proceed to Step 6 to solve for self-accommodated martensite. If these conditions are not satisfied, the step is not fully elastic and the code proceeds sequentially through this algorithm to solve for the transformation strain.

4. If \( F_{\sigma-re} \leq 0 \), skip to Step 5, else the transformation strain is reoriented by solving the reorientation residuals for \( \sigma, \Delta \varepsilon_{re}, \) and \( \Delta \varepsilon_{eq} \):

\[
0 = \sigma - C : (e - e^n - \Delta \varepsilon_{re})
\]

(23)

\[
0 = \Delta \varepsilon_{re} - \Delta \varepsilon_{eq}(e^n) : \sigma'
\]

(24)

\[
0 = ||\tilde{\tau}(e^n) : \sigma'|- Y_{re}|
\]

(25)

Note that Eq. (24) could be substituted into Eq. (23) to further reduce the dimension of the problem, but because we wish to output \( \Delta \varepsilon_{eq} \), we choose to program this less compact system of residuals. If \( F_{\sigma-re} \leq 0 \), skip to Step 6, else the transformation strain is evolved by solving the orientation residuals for \( \sigma, \Delta \varepsilon_{eq}, \tilde{\sigma}, \), and \( \Delta \varepsilon_{eq} \):

\[
0 = \sigma - C : (e - e^n - \Delta \varepsilon_{eq})
\]

(26)

\[
0 = \Delta \varepsilon_{eq} - \Delta \varepsilon_{eq}(e^n) : \sigma'
\]

(27)

\[
0 = ||\tilde{\tau}(e^n) : \sigma'|- Y_{re}|
\]

(28)

\[
\tilde{\sigma} - \frac{||\sigma'||}{\sqrt{3/2}e^n_{\text{max}}}
\]

(29)

where \( X_\sigma \) follows the definition given in Eq. (10). Akin to Step 4, Eq. (27) could be substituted into Eq. (26) if the implementer did not care to track the evolution of \( \Delta \varepsilon_{eq} \), but because we do choose this less compact system of residuals. Also note that we include \( \Delta \varepsilon_{eq} \) in Step 4 of Eq. (26), thus if \( \varepsilon_{eq} \) evolved in Step 4, \( \sigma \) will be calculated in Step 5 for the total transformation strain.

5. In the case of initiation of transformation strain (i.e., while \( ||\sigma'|| \leq \tilde{\sigma} \)), \( N(e^n) \) is numerically singular and the initiation residuals are used (we refer the reader to [1] for their derivation):

\[
0 = \sigma - C : (e - \Delta \varepsilon_{eq})
\]

(30)

\[
0 = \Delta \varepsilon_{eq} - \Delta \varepsilon_{eq}(e^n) \sigma'
\]

(31)

\[
0 = ||\tilde{\tau}(e^n) : \sigma'|- Y_{re}|
\]

(32)

\[
	ilde{\sigma} - \frac{||\sigma'||}{\sqrt{3/2}e^n_{\text{max}}}
\]

(33)

where the deviatoric stress is used to define the initial transformation strain direction:

\[
X_\sigma = \sigma' - \frac{\rho(\bar{T}^0 - \bar{u}_0) + H_{\text{tr}} e^n}{\sqrt{3/2}e^n_{\text{max}}} N(\sigma')
\]

(34)

6. Finally, the evolution of self-accommodated martensite is determined. The self-accommodation driving force is evaluated:

\[
X_{SA} = -\rho(\bar{T}^0 - \bar{u}_0) = -\rho \bar{u}_0 \left( T - M^n_0 \right)
\]

(35)

The self-accommodation limit function is then calculated:

\[
F_{SA} = \begin{cases} 
X_{SA} - Y_{SA}(\zeta_{SA}^n) : X_{SA}\zeta_{SA}^n > 0 \\
- X_{SA} - Y_{SA}(\zeta_{SA}^n) : X_{SA}\zeta_{SA}^n < 0
\end{cases}
\]

(36)

using the kinetic equations:

\[
Y_{SA}' = \rho \zeta_{SA}^n = \rho \bar{u}_0 \left( M^n_0 - M^n_0 \right) \zeta_{SA}^n
\]

(37)

\[
Y_{SA}' = \rho \zeta_{SA}^n (1 - \zeta_{SA}^n) + Y_{SA}^\text{max} + \sigma_{SA}^\text{max}
\]

(38a)

\[
= \rho \bar{u}_0 \left( \zeta_{SA}^n M^n_0 + (A^n_j - A^n_{ij}) (1 - \zeta_{SA}^n) + \frac{\sigma}{C_M} \right)
\]

(38b)

If the condition

\[
F_{SA} \leq 0
\]

(39)
is satisfied, self-accommodated martensite does not evolve. Otherwise, \( \Delta \xi_{SA} \) is incremented through:

\[
\Delta \xi_{SA} = \frac{\Delta \xi}{\xi_{SA}} + \Delta \xi_{SA}
\]  

(40)

where to impose the bounds on the volume fractions (Eq. (5)), if applicable, the calculation of \( \Delta \xi_{SA} \) is made according to:

\[
\Delta \xi_{SA} = -\Delta \xi_{\sigma} \text{ if } \frac{\Delta \xi}{\xi_{SA}} + \frac{\Delta \xi_{\sigma}}{\xi_{SB}} > \xi_{\text{max}} \text{ or } T < M^0 \text{ and } \frac{\xi_{\sigma}}{\xi_{SB}} + \Delta \xi_{\sigma} < \xi_{\text{max}}
\]  

(41)

otherwise:

\[
\Delta \xi_{SA} = \begin{cases} 
-\frac{\mu p}{\rho} \Delta T & \text{if } F_{SA}(T, \xi_{SB}) > 0 \text{ and } \Delta T < 0 \\
-\frac{\mu p}{\rho} \Delta T + \frac{\xi_{\sigma}}{\xi_{SB}} + \Delta \xi_{\sigma} & \text{if } F_{SA}(T, \xi_{SB}) < 0 \text{ and } \Delta T > 0 
\end{cases}
\]  

(42)

7. The final calculated stress is passed back to Abaqus, along with the user-defined state variables transformation strain, orientation and reorientation strain increments, and martensite volume fractions, return to Step 1.

2.2. Discussion of improvements

2.2.1. Numerically robust kinetic law

The most significant modification for robust numerical implementation is assumption of a new kinetic law for the frictional force resisting orientation (Eq. (20)). For discussion purposes, the originally proposed kinetic equations governing orientation follow:

\[
Y_{\sigma}(\xi_{\sigma}) = \begin{cases} 
A' \xi_{\sigma} - B' \xi_{\sigma} \ln(1 - \xi_{\sigma}) + C' & \text{if } \xi_{\sigma} > 0 \\
A'(1 - \xi_{\sigma}) - B'(1 - \xi_{\sigma}) \ln(\xi_{\sigma}) + C' & \text{if } \xi_{\sigma} < 0
\end{cases}
\]  

(43)

The value of this function represents a critical effective stress that must be overcome for the orientation mechanism of the model to activate. While in theory the natural logarithm terms drive the value of \( Y_{\sigma} \) to \( \infty \) as the volume fraction saturates (0 \( \leq \xi_{\sigma} \leq 1 \)), in practice this is not the case.

Discontinuities would arise in these equations if \( \xi_{\sigma} = 0 \) or 1, thus numerically \( \xi_{\text{min}} \) a small number very close to 0, must be chosen as the lower bound, and \( \xi_{\text{max}} = 1 - \xi_{\text{min}} \) is a logical corresponding choice for an upper bound. Table 1 and Fig. 3 show that in decreasing \( \xi_{\text{min}} \) by orders of magnitude toward the limit of machine precision, the logarithm terms approach \( 0 \). This does not numerically replicate the intended limit of \( \infty \). The orientation mechanism is controlled by the limit function Eq. (18) and in the state of saturated volume fraction, \( \xi_{\text{SA}} \) varies proportionally with \( \sigma^* \) (see Eqs. (10) and (34)) The result using Eq. (43) in Eq. (18) is that if the applied load is increased a modest amount after the saturation of oriented martensite, situation may arise where \( F_{\sigma} > 0 \) even though it is not physically possible for martensite to further orient. For example, consider the case in which orientation of martensite during forward transformation saturates at an effective stress of 500 MPa. If we chose \( \xi_{\text{min}} = 10^{-12} \), \( B' = 5 \text{ MPa, } A' = 0 \text{ MPa, and } C' = 70 \text{ MPa} \), all reasonable choices for a polycrystalline SMA such as NiTi, \( \eta_{SB} = 138 \text{ MPa when } \xi_{\sigma} = 1 \). Thus if we continue to load the material to the vicinity of 700 MPa, \( F_{\sigma} > 0 \) and the orientation mechanism will attempt to increase the transformation strain, even though physically this is impossible for the material given our constitutive assumptions (Eqs. (4)–(6)). Thus, we conclude that the form of a soft constraint built into this kinetic law is too soft for robust numerical implementation.

To make this kinetic law more numerically robust and also to reduce the number of calibration constants needed for the model, we have replaced the original assumed form (Eq. (20)) with Eq. (43). In choosing the tangent function instead of the natural logarithm we introduce a soft constraint that provides sufficient hardening to numerically bound \( \xi \); the constraint continues to grow by an order of magnitude in each order of magnitude toward saturation of \( \xi \) (Table 1). Additionally, the need for two calibration constants \( B' \) and \( B' \) is eliminated. In the original model these constants were required to impose a constraint to bound \( \xi \), and could be adjusted to control the “roundness” of the exit side of the hysteresis loop in each forward and reverse transformation. Since the bound is now sufficiently hard, a scaling constant is no longer required. Since the motivation for the model is to replicate the mechanics of the overall transformation strain, not smoothing of the hysteresis corners, it is not necessary to reintroduce a tuning parameter for hysteresis smoothing. Furthermore, the constant \( A' \) was redundant considering \( H_{\sigma} \) (see Eqs. (17), (18) and (43)), thus it has also been removed in this improved kinetic law. \( A' \) remains as a tuning parameter that allows the modeler to calibrate a unique hardening slope for reverse transformation with respect to forward transformation.

2.2.2. Reduced set of model inputs easily interpreted from empirical measurements

In [1], model inputs were defined through 17 scalars: the 6 phenomenological material properties \( E, v, \rho, \tilde{u}_0, \tilde{n}_0, \xi_{\text{max}} \) combined with 3 model material parameters \( c', c'', Y_{TB} \) and 8 kinetic tuning parameters \( H_0, Y_{SB}, A', B', C', A', C', C' \). However, it is also possible to instead write the parameters \( c', c'', Y_{TB} \) along with the material properties \( \tilde{u}_0, \tilde{n}_0 \) in terms of the four stress-free transformation

<table>
<thead>
<tr>
<th>( \xi_{\text{min}} )</th>
<th>( -\ln (\xi_{\text{min}}) )</th>
<th>( \tan (\pi (\xi_{\text{min}} + \frac{1}{2}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00E-02</td>
<td>4.61E+00</td>
<td>3.18E+01</td>
</tr>
<tr>
<td>1.00E-03</td>
<td>6.91E+00</td>
<td>3.18E+02</td>
</tr>
<tr>
<td>1.00E-04</td>
<td>9.21E+00</td>
<td>3.18E+03</td>
</tr>
<tr>
<td>1.00E-05</td>
<td>1.15E+01</td>
<td>3.18E+04</td>
</tr>
<tr>
<td>1.00E-06</td>
<td>1.38E+01</td>
<td>3.18E+05</td>
</tr>
<tr>
<td>1.00E-07</td>
<td>1.61E+01</td>
<td>3.18E+06</td>
</tr>
<tr>
<td>1.00E-08</td>
<td>1.84E+01</td>
<td>3.18E+07</td>
</tr>
<tr>
<td>1.00E-09</td>
<td>2.07E+01</td>
<td>3.18E+08</td>
</tr>
<tr>
<td>1.00E-10</td>
<td>2.30E+01</td>
<td>3.18E+09</td>
</tr>
<tr>
<td>1.00E-11</td>
<td>2.53E+01</td>
<td>3.18E+10</td>
</tr>
<tr>
<td>1.00E-12</td>
<td>2.76E+01</td>
<td>3.18E+11</td>
</tr>
<tr>
<td>1.00E-13</td>
<td>2.99E+01</td>
<td>3.18E+12</td>
</tr>
<tr>
<td>1.00E-14</td>
<td>3.22E+01</td>
<td>3.18E+13</td>
</tr>
<tr>
<td>1.00E-15</td>
<td>3.45E+01</td>
<td>3.18E+14</td>
</tr>
</tbody>
</table>
temperatures and the Patel–Cohen coefficient. This formulation is desirable as the specific internal energies and entropies associated with the austenite–martensite phase transformation are only occasionally reported, but stress-free transformation temperatures are almost always documented. Additionally, in numerically improving the orientation kinetic law (Eq. (20), discussed further in Section 2.2.1) the number of kinetic tuning parameters was reduced from seven to four. This results in a set of 14 model inputs consisting of 9 material properties that are readily published for nearly all SMAs. In addition, 2 martensite transformation kinetic parameters, 2 for each variant, and 3 kinetic tuning parameters are taken from empirical measurements. Since phase transformation and orientation are driven by the deviatoric stress state introduced the form of Eq. (12), as this new form allows the kinetic parameters to be directly calculated for convenience of shorthand within the numerical implementation:

\[
\eta_0 = \frac{C_M k_{\text{max}}}{\rho}
\]  
(44)

\[
\tilde{u}_0 = \tilde{u}_0 M_i^0
\]  
(45)

\[
c^f = \tilde{u}_0 (M_i^0 - M_i^f)
\]  
(46)

\[
c^r = \tilde{u}_0 (A_i^0 - A_i^r)
\]  
(47)

\[
Y_{\text{tr}} = \rho \tilde{u}_0 (A_i^0 - M_i^f).
\]  
(48)

One final modification may be made to the model inputs to facilitate direct translation from empirical measurements. Since phase transformation and orientation are driven by the deviatoric stress in this model, the kinetic parameters taken from empirical measurements should be scaled by \( \sqrt{2/3} \) prior to calculations. This may either be done manually, or else written into the numerical code.

2.2.3. New choice of reorientation limit function

The original proposed reorientation limit function was

\[
F_{\text{re}} = \frac{1}{2} \sigma : \tilde{I} e_{\text{re}} : \sigma - Y_{\text{re}}
\]  
[1]. In this improved model, we have instead introduced the form of Eq. (12), as this new form allows the dissipation of reorientation \( Y_{\text{re}} \) to be expressed in units of stress rather than stress squared. Hence, it is now possible to calibrate this parameter through phenomenological observation of an effective onset stress for reorientation, just as an effective onset stress for orientation is used to calibrate \( Y_{\text{re}} \). Note that \( \| \tilde{I} e_{\text{re}} \| : \sigma \) and \( \sqrt{\sigma : \tilde{I} e_{\text{re}} : \sigma} \) are equivalent expressions in choosing to implement either form of \( \tilde{I} e_{\text{re}} \) (see Section 2.1, Step 3) and \( t(r e_{\text{re}}) = 0 \).

2.2.4. Correction to the specific Entropy–Patel–Cohen relation

In the original work, Panico and Brinson defined the Patel–Cohen coefficient relationship to the specific entropy as \( C_M = \rho \Delta U_0 \) [1]. However, by definition the Patel–Cohen relationship between stress and energy of transformation scales as the product of stress and transformation strain [31]. The omission of transformation strain in this relationship resulted in an issue of improper scale. In the improved model, Eq. (44) was adopted from [32] to correct the calculations. Note that for purposes of continuum formulation, we have oversimplified the original description of this relationship, which considered microstructure at the level of habit planes formed by pairs of variants [31].

3. Numerical schemes for solving orientation and reorientation equations

There are four scenarios in the model that require evolution of the internal variables. \( \varepsilon_{\sigma} \) evolves when Eq. (39) is not satisfied. However, the evolution of self-accommodated martensite may be calculated using the derived explicit solution procedure (Step 6, Eqs. 35–42), thus it requires no further numerical consideration. Hence there are three scenarios that require sets of coupled equations to be solved:

(i) When Eq. (13) is not satisfied for \( F_{\sigma} \) and \( \| e_\sigma \| \leq \varepsilon, \varepsilon_\sigma \) and \( e_\sigma \) will initiate via Eqs. 30–34;

(ii) When Eq. (13) is not satisfied for \( F_{\sigma} \), \( \varepsilon_\sigma \) will orient via Eqs. 23–25;

(iii) When Eq. (13) is not satisfied for \( F_{\sigma} \) and \( \| e_\sigma \| > \varepsilon, \varepsilon_\sigma \) will orient via Eqs. 26–29.

As mentioned in the introduction, two numerical schemes for solving these sets of equations are presented in this section. In Section 4, they are compared through studies of convergence and computational efficiency.

3.1. Scheme 1-constant lagrange multiplier increment

This is the simplest of the schemes to implement and the scheme that was used to augment the implicit integration scheme initially proposed by Panico and Brinson (see Section 1.2). A fixed Lagrange multiplier increment is used for both orientation and reorientation \( \Delta e_{\sigma,\text{re}} = \alpha_0 \). When orientation and/or reorientation Kuhn–Tucker conditions require inelastic deformation to be satisfied (i.e., Eq. (13)), this constant multiplier is used in the strain rate equations (Eq. (11)), and the new transformation strain (Eq. (2)) is calculated along with the updated stress (Eq. (26)) and martensite volume fractions (Eq. (29) or (33), depending on Scenario (i) or (iii) stated above). The error from not solving the equations through integration propagates into the next step and is compensated for in subsequent calculations. That is, at each step of the finite element simulation, Eq. (13) will continue to not be satisfied until the solution to the equations of the active scenario have converged. While this scheme is simple and the easiest scheme to program, the choice of \( \alpha_0 \) relative to the parameters of the finite element simulations (mesh, time) is sensitive, as demonstrated in Section 4.

With this scheme, the algorithm for Steps 4 and 5 is:

- \( \Delta e_{\sigma,\text{re}} = \alpha_0 = \text{constant} \)
- Plug \( \Delta e_{\sigma,\text{re}} \) into Eqs. (24) and (27) and solve for \( \Delta \varepsilon_{\sigma,\text{re}} \)
- Use to update \( \sigma, \varepsilon_\sigma \) through Eqs. (2), (26), (29)

END

3.2. Scheme 2-explicit integration

A new explicit scheme was developed to relieve sensitivity inherent to the “Constant Lagrange Multiplier” scheme while avoiding the cumbersome implementation of the Jacobian for the “Implicit Integration” scheme (see Appendix B). This new integration scheme is conceptually simple, analogous to the idea of a radial return algorithm [33]. However, for this specific constitutive model, robust implementation requires a bit of additional derivation. The gist is to begin with an initial guess of the Lagrange multiplier increments, and then continue to increase \( \Delta \varepsilon_{\sigma,\text{re}} \) until the solution converges.

First, we introduce a new computation input \( \zeta \), the targeted number of iterations for each integration. The choice of \( \zeta \) is studied in Section 4. However, generally, choosing smaller \( \zeta \) leads to reduced computation time at the expense of less precise convergence while larger \( \zeta \) has opposite consequences.

The derivation commences with the observation that each mechanism will require integration if and only if Eq. (13) is not satisfied. Consequently, in solving for the new transformation strain, stress, and martensite volume fractions, the algorithm will loop until Eq. (13) is satisfied. This in mind, when Eq. (13) is not
satisfied, we would like each step of the explicit integration scheme to increment the value of the limit function by
\[ \Delta F_{\text{tr}, \text{re}} = -\frac{F_{\text{tr}} \{ \sigma^{\text{trial}} \}}{\zeta} \] (49)

until we converge to a solution that satisfies Eq. (13). Examining the orientation Eqs. (10) and (12) in rate form and noting that \( \mathbf{N}(e_{\text{t}}) \approx 0 \) since the transformation strain is always normalized by its magnitude, and \( Y_{\text{e}} \approx 0 \) as well since we are explicitly stepping through time in small increments:
\[ \dot{F}_{\text{tr}} = \frac{X_{\sigma} : X_{\sigma}}{\|X_{\sigma}\|} - Y_{\alpha} \approx \dot{\sigma} : X_{\sigma} \] (50)

To study the order of magnitude of the reorientation limit function in Eq. (12), it is useful write the expressions \( \mathbf{N}(e_{\text{t}}) : \sigma \) and \( \| \mathbf{N}(e_{\text{t}}) : \sigma \| \) out in indical notation and simplify using \( \text{tr}(e_{\text{t}}) = 0 \). This calculation results in:\[ \mathbf{N}(e_{\text{t}}) : \sigma = \sigma - e_{\text{t}} : \sigma_{\text{tr}} : e_{\text{t}} \] (51)

\[ \| \mathbf{N}(e_{\text{t}}) : \sigma \| = \sqrt{\sigma : \sigma - \frac{(e_{\text{t}} : \sigma)^2}{e_{\text{t}} : e_{\text{t}}}} \] (52)

Making the appropriate substitution and taking the time derivative results in:
\[ \dot{F}_{\text{tr}} = \frac{\sigma : \sigma - \frac{1}{2} \text{tr}(\sigma) \text{tr}(\sigma) - (e_{\text{t}} : \sigma + e_{\text{t}} : \sigma_{\text{tr}} : e_{\text{t}} : \sigma_{\text{tr}} : e_{\text{t}}) e_{\text{t}} : e_{\text{t}}}{\mathbf{N}(e_{\text{t}}) : \sigma} \] (53)

Next, we recognize that we are solving the reorientation and orientation equations individually, and while they are being solved, the total strain increment is fixed and we are beginning from the assumption that the total strain increment was elastic (Section 2.1, Steps 1–3). Thus, in isolation the stress rates during the integration of the equations for each mechanism are (Section 2.1, Steps 4, 5):
\[ \dot{\sigma} = -C : \dot{e}_{\text{t}} \quad \text{and} \quad \dot{\sigma} = -C : \dot{e}_{\text{t}}. \] (54)

Substituting these expressions into Eqs. (50) and (53), and also Eq. (27) into Eqs. (50) and (24) into Eq. (53) (all in rate form), we arrive at the expressions:
\[ \dot{F}_{\text{tr}} = -\int \left( C : \dot{e}_{\text{t}} \right) : X_{\sigma} \approx -\dot{e}_{\text{t}} : X_{\sigma}, \] (55)

\[ \dot{F}_{\text{tr}} = \frac{\left( \sigma - e_{\text{t}} : \sigma_{\text{tr}} : e_{\text{t}} : \sigma_{\text{tr}} : e_{\text{t}} \right) e_{\text{t}} : e_{\text{t}}}{\mathbf{N}(e_{\text{t}}) : \sigma} \] (56)

Finally, these may be substituted into Eq. (49) in incremental form to derive the initial values of the Lagrange multiplier increments for this explicit integration scheme:
\[ \Delta \lambda_{\sigma, \text{tr}}^{0} = \frac{\|X_{\sigma} - \mathbf{C} \cdot X_{\sigma}\|}{X_{\sigma} : \mathbf{C} : X_{\sigma}} \] (57)

\[ \Delta \lambda_{\text{tr}, \text{re}}^{0} = \frac{\| \sigma - e_{\text{t}} : \sigma_{\text{tr}} : e_{\text{t}} : \sigma_{\text{tr}} : e_{\text{t}} \|}{\mathbf{N}(e_{\text{t}}) : \sigma} \] (58)

Using these initial values, the equations for orientation and reorientation (Steps 4 & 5) may now be independently solved through the explicit iteration scheme:

\[ \Delta \lambda_{\sigma, \text{tr}}^{0} = \Delta \lambda_{\sigma, \text{tr}}^{0} \]
\[ \# \text{ iterations} = 1 \]

While Eq. (13) not satisfied and \# iterations \( \leq \) max iterations
\[ \circ \text{ Substitute } \Delta \lambda_{\sigma, \text{tr}}^{0} \text{ into Eqs. (24) and (27) and solve for } \Delta \lambda_{\text{tr}, \text{re}} \]
\[ \circ \text{ Use to update } \sigma, \zeta_{\text{tr}, \text{re}}, \dot{X}_{\sigma} \text{ through Eqs. (2), (18), (19), (26), (29)} \]
\[ \circ \# \text{ iterations} = \# \text{ iterations} + 1 \]
\[ \Delta \lambda_{\text{tr}, \text{re}} = \# \text{ iterations} \times \Delta \lambda_{\text{tr}, \text{re}}^{0} \]

END

These multipliers cannot exactly converge in a single step \( (\zeta = 1) \) for all cases since there was some approximation made in their derivations (assumptions that \( Y_{\sigma}, \mathbf{N}(e_{\text{t}}) \approx 0 \)) and because they are calculated by reducing tensors to scalars, and then multiplying those scalars by components of tensors again. For the same reasons, we cannot expect to always converge in exactly the targeted number of iterations \( \zeta > 1 \). In the current work we choose max iterations = 10, and we will document the actual number of iterations used for different choices of \( \zeta \) in Section 4.

Checks may be implemented to understand if the choices of \( \zeta \) and max iterations are appropriate: (1) warning messages may be printed to the log file if the procedure stopped due to maximum # iterations; (2) in the case of reorientation, numerical implementation of the tangency condition (Eq. (7))
\[ e_{\text{t}} : e_{\text{r}} < \zeta \]

may be used to check sensibility of the calculated reorientation strain increment (the same effective small strain limit \( \zeta \) as is used to protect against singularity of \( \mathbf{N}(e_{\text{t}}) \) may be used here). If there are problems due to poor choice of material parameters, simulations parameters, mesh, boundary conditions, or bugs in the code, one of these two conditions will not be met fairly quickly in the simulation.

4. Results and discussion of simulations using the explicit numerical schemes

Having formulated two schemes for explicit integration of the improved constitutive model within an explicit finite element framework, we proceed to study convergence of each scheme with respect to the computational parameters \( \lambda_{\sigma} \) for Scheme 1 (Section 3.1) and \( \zeta \) for the Scheme 2 (Section 3.2). Also, since this rate independent constitutive model is being implemented into an explicit FE framework, it is desirable to have a numerical implementation that is insensitive to liberal mass scaling. Mass scaling artificially inflates the mass of the material to force Abaqus to calculate a larger stable time increment for the simulation. In a quasi-static simulation using explicit FE, such as one using a rate-independent material model, mass scaling may be used to reduce computation time (larger time increment = fewer simulation steps = less computation time, provided the order of convergence

\[ \Delta \lambda_{\sigma, \text{tr}} = \Delta \lambda_{\sigma, \text{tr}}^{0} \]
\[ \# \text{ iterations} = 1 \]

While Eq. (13) not satisfied and \# iterations \( \leq \) max iterations
\[ \circ \text{ Substitute } \Delta \lambda_{\sigma, \text{tr}}^{0} \text{ into Eqs. (24) and (27) and solve for } \Delta \lambda_{\text{tr}, \text{re}} \]
\[ \circ \text{ Use to update } \sigma, \zeta_{\text{tr}, \text{re}}, \dot{X}_{\sigma} \text{ through Eqs. (2), (18), (19), (26), (29)} \]
\[ \circ \# \text{ iterations} = \# \text{ iterations} + 1 \]
\[ \Delta \lambda_{\text{tr}, \text{re}} = \# \text{ iterations} \times \Delta \lambda_{\text{tr}, \text{re}}^{0} \]

END

It may also be desirable to use these substitutions in coding the model as they allow omission of explicit calculation of the projection tensor.
4.1. Simulations of a single 3D brick element

A single 8-node explicit linear brick element (type C3D8R [8]) was subjected to an axial-shear, non-proportional, displacement-control loading path. These loadings were imposed by fixing the bottom surface of the element in the 1- and 2-directions and then displacing the nodes of the top surface of the element as indicated in Table 3 and Fig. 5. The converged axial and shear components of the calculated responses are shown in Figs. 5 and 6, along with the equivalent Cauchy stress and Lagrangian strain responses calculated via

$$\sigma_{eq} = \sqrt{\sigma_{22}^2 + 3\sigma_{12}^2}, \quad \varepsilon_{eq} = \sqrt{\varepsilon_{22}^2 + \frac{4}{3}\varepsilon_{12}^2}. \quad (60)$$

The converged oriented martensite volume fraction evolution is also shown in Fig. 5c. In examining Figs. 5b and 6b, the stress response may seem counter-intuitive as there are many inflection points that do not correspond to inflection points in the applied loading path. These inflection points are a result of the driving forces and kinetics of orientation and reorientation (Eqs. (10), (18), (19), (20)) – because of hysteresis as well as saturation events, the initiation and cessation of activity of these mechanisms does not necessarily coincide with inflections in the mechanical loading of the material. A more detailed discussion of the mechanics of this converged response is given in Appendix C.

4.1.1. Scheme 1 convergence sans mass scaling

The responses of simulations using Scheme 1 while varying \(\lambda_0\) in order of magnitude increments from \(10^{-5}\) to \(10^{-12}\) are shown in stress-space through Fig. 7a and b. In examining the responses on the scale of Fig. 7a it is apparent that for \(\lambda_0\) varying from \(10^{-10}\) to \(10^{-12}\) stresses are grossly over-predicted. Because the value of \(\lambda_0\) is not large enough, the explicit finite element continues to calculate an elastic response using a modulus of 70 GPa in the absence of the VUMAT calculating a reasonable estimation of the inelastic transformation strain. In examining these same responses on the scale of Fig. 7b, for \(\lambda_0\) varying from \(10^{-6}\) to \(10^{-5}\) the stress predictions become non-convergent, over and undershooting the solution as \(\lambda_0\) is now too large. Thus, for this material parameterization, element choice, and prescribed loading, it is demonstrated that selecting \(\lambda_0\) between \(10^{-7}\) and \(10^{-9}\) results in convergence of the mechanical response. Figs. 7c & d are used to compare Schemes 1 & 2, and will be discussed in Section 4.1.3.

These results demonstrate the ability of this scheme to simulate the thermodynamics of the constitutive model. Furthermore, the explicit finite element calculation is able to compensate for a fairly wide range of \(\lambda_0\) choices in the VUMAT calculation, spanning two orders of magnitude. One drawback of this scheme is that the convergence study must be run to understand if the choice of \(\lambda_0\) is reasonable – there was no way for us to know a priori that \(\lambda_0\) between \(10^{-7}\) and \(10^{-9}\) would converge; we had to employ the guess-and-check method. For more complex models, this convergence study becomes quite tedious when checking many integration points. Further insights in choice of this scheme will be made in comparing this scheme with Scheme 2 in Section 4.1.3.

4.1.2. Scheme 2 convergence sans mass scaling

Mechanical response predictions (stress, strain, martensite volume fraction) converged to those shown in Figs. 5 and 6 for choices of \(\zeta\) ranging from 1 to 500. Insight as to why may be gained considering Fig. 8 in the context of the convergence study of the previous section. It was demonstrated that the finite element calculation could accommodate two orders of magnitude difference in the Lagrange multiplier increment used in the VUMAT. Fig. 8 shows that by using Eqs. (57) and (58) to calculate the initial Lagrange multiplier increments and Scheme 2 to integrate the sets of transformation and/or reorientation equations for Steps 4 & 5, the calculated initial Lagrange multiplier increments always fall within an

Table 2
Material properties and kinetic parameters used in the simulations.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>70 GPa</td>
<td>(\nu_0)</td>
<td>40 MPa</td>
</tr>
<tr>
<td>(v)</td>
<td>0.4</td>
<td>(H_0)</td>
<td>200 J/kg</td>
</tr>
<tr>
<td>(\rho)</td>
<td>6.45 g/cm³</td>
<td>(C)</td>
<td>100 MPa</td>
</tr>
<tr>
<td>(M_s)</td>
<td>258 K</td>
<td>(C_f)</td>
<td>100 MPa</td>
</tr>
<tr>
<td>(M_l)</td>
<td>248 K</td>
<td>(A_r)</td>
<td>0 MPa</td>
</tr>
<tr>
<td>(A_s)</td>
<td>278 K</td>
<td>(A_f)</td>
<td>288 K</td>
</tr>
<tr>
<td>(C_{AM})</td>
<td>5.5 MPa/K</td>
<td>(e_{\text{min}})</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Fig. 4. The brick element and the defined coordinate system.

Table 3
Simulated displacement-control loadings applied to the 4 top nodes of the simulated element. All displacements were applied with a velocity of 0.16 mm/s.

<table>
<thead>
<tr>
<th>Step</th>
<th>Time</th>
<th>Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>+y</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>+x</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>–y</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>–x</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>+y</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>+x</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>–y</td>
</tr>
</tbody>
</table>
order of magnitude of the converged solution (the solution converged at choices of \( \zeta = 100, 500 \)). Considering the case \( \zeta = 1 \), the lack of convergence of the Lagrange multipliers but convergence of the stress, strain, and volume fraction calculations in Figs. 5 and 6 is significant. It means that using Scheme 2, one may choose not to integrate the coupled PDEs in Steps 4 & 5 at all (i.e., \( \zeta = 1 \)).
and still simulate the correct stress, strain, and volume fraction responses by allowing Abaqus Explicit FE calculations to compensate for any error caused by the lack of exactness in calculating the transformation/reorientation Lagrange multipliers. Furthermore, choosing a small number of iterations (i.e., \( f = 5 \)) quickly allows the solution to approach the converged solution for all internal and external variables, and very little difference was seen between the simulated Lagrange multiplier increments and the converged solution in choosing \( f \) greater than 50 (see Fig. 8c).

Another interesting observation is the number of iterations used in integrating the equations for Steps 4 & 5, plotted in Fig. 9. Recall from Section 3.2 that the targeted number of iterations \( \zeta \) could only be treated as an approximation since tensors were being contracted to scalars in Eqs. (57) and (58). Thus, we chose to allow for an order of magnitude more iterations in the VUMAT by setting max iterations = \( 10^f \). However, Fig. 9 shows that while \( \zeta \) was an approximation, there were never more than \( f \) iterations required for convergence, regardless of the choice of \( \zeta \). Thus, an implementer of Scheme 2 could restrict the maximum number of iterations to \( \zeta + 1 \) instead of \( 10^f \), though we note that this result does not automatically transcend to multiple element meshes subjected to non-proportional loading.

Fig. 7. Stress responses of simulations using the Constant Lagrange Multiplier scheme (see Section 3.1). (a) and (b) depict the responses of varying \( \lambda_0 \) sans mass scaling; (c) and (d) show the result of varying the mass scaling factor for \( \lambda_0 = 10^{-8} \).

Fig. 8. Calculations of (a) the orientation strain Lagrange multiplier increments and (b) the reorientation strain Lagrange multiplier increments during simulations using the Explicit Integration scheme (see Section 3.2) as \( \zeta \) was varied. (c) is a magnified view of the reorientation multiplier calculations about \( t = 70 \), showing that while the \( \zeta = 50 \) simulation is indistinguishable from the converged solution on the scale of figures (a) and (b), it is not fully converged when examined more closely.

Fig. 9. Recall from Section 3.2 that the targeted number of iterations \( \zeta \) could only be treated as an approximation since tensors were being contracted to scalars in Eqs. (57) and (58). Thus, we chose to allow for an order of magnitude more iterations in the VUMAT by setting max iterations = \( 10^\zeta \). However, Fig. 9 shows that while \( \zeta \) was an approximation, there were never more than \( \zeta \) iterations required for convergence, regardless of the choice of \( \zeta \). Thus, an implementer of Scheme 2 could restrict the maximum number of iterations to \( \zeta + 1 \) instead of \( 10^\zeta \), though we note that this result does not automatically transcend to multiple element meshes subjected to non-proportional loading.
4.1.3. Comparison of the schemes

Through consideration of the previously presented results (i.e., Sections 4.1.1 and 4.1.2), the most significant difference between the two schemes is that Scheme 1 must be manually calibrated and studied for convergence when mesh and/or model parameters are changed while Scheme 2 is self-calibrating, even when explicit integration of the PDEs for orientation and reorientation is turned off (i.e., \( \zeta \) set to 1). The question for a user of this constitutive model then becomes “Does the coding complexity and computational cost of calculating Eqs. (57) and (58) outweigh the benefits of self-calibration?” While the final answer to this question may be subject to both personal opinion and frequency of use of the model, in the remainder of this section we will quantify the computational cost of each scheme both with and without mass scaling for a single 3D element, and in Section 4.2 this study of numerical simplicity of Scheme 1 vs. benefits of self-calibration of Scheme 2 will be extended through simulations of a more complicated mesh.

To aid in the ensuing discussion, CPU times for the simulations both with and without mass scaling are listed in Table 4. These simulations were performed on a single core of a Dell T5500 Linux workstation. For the ensuing mass scaling study, \( k_0 = 10^8/C_0 \) was chosen for Scheme 1 since it was in the middle of the observed convergence interval of \( 10^{-7} \leq k_0 \leq 10^{-9} \) while Scheme 2 was studied for both no integration, i.e., \( \zeta = 1 \), and \( \zeta = 50 \). Mass scaling was imposed by use of a mass scaling factor \( [8] \), which was varied in order of magnitude from 1 (no mass scaling) to \( 10^8 \).

Comparing the results for no mass scaling, Scheme 2 without integration (\( \zeta = 1 \)) took 3% longer to run. Both schemes provided converged stress, strain, and volume fraction calculations, but neither scheme provided a converged Lagrange multiplier solution (see the preceding two sections and Figs. 5–8). However, the Scheme 2 Lagrange multiplier calculations for \( \zeta = 1 \) were of the same order of magnitude as the converged solution (Fig. 8), while the Scheme 1 fixed Lagrange multiplier increment was off by an order of magnitude. Integration of the PDEs using \( \sim 50 \) iterations increased the time by an additional 6%, while providing the additional accuracy of a converged Lagrange multiplier solution. Further increase of the number of iterations for each integration continued to increase the computational cost with virtually no benefit, as the Lagrange multiplier calculations did not experience substantial improvement (Fig. 8c).

To aid in the ensuing discussion, CPU times for the simulations both with and without mass scaling are listed in Table 4. These simulations were performed on a single core of a Dell T5500 Linux workstation. For the ensuing mass scaling study, \( k_0 = 10^8/C_0 \) was chosen for Scheme 1 since it was in the middle of the observed convergence interval of \( 10^{-7} \leq k_0 \leq 10^{-9} \) while Scheme 2 was studied for both no integration, i.e., \( \zeta = 1 \), and \( \zeta = 50 \). Mass scaling was imposed by use of a mass scaling factor \( [8] \), which was varied in order of magnitude from 1 (no mass scaling) to \( 10^8 \).

Comparing the results for no mass scaling, Scheme 2 without integration (\( \zeta = 1 \)) took 3% longer to run. Both schemes provided converged stress, strain, and volume fraction calculations, but neither scheme provided a converged Lagrange multiplier solution (see the preceding two sections and Figs. 5–8). However, the Scheme 2 Lagrange multiplier calculations for \( \zeta = 1 \) were of the same order of magnitude as the converged solution (Fig. 8), while the Scheme 1 fixed Lagrange multiplier increment was off by an order of magnitude. Integration of the PDEs using \( \sim 50 \) iterations increased the time by an additional 6%, while providing the additional accuracy of a converged Lagrange multiplier solution. Further increase of the number of iterations for each integration continued to increase the computational cost with virtually no benefit, as the Lagrange multiplier calculations did not experience substantial improvement (Fig. 8c).

As mentioned in the introduction to Section 4, since this constitutive model is rate independent, it is desirable to have a numerical implementation that is insensitive to mass scaling so this technique may be invoked to reduce the computational cost of simulations. Stress responses of Scheme 1 as the mass scaling factor is varied

![Table 4](image)

**Table 4**

<table>
<thead>
<tr>
<th>Factor</th>
<th>None</th>
<th>10</th>
<th>10²</th>
<th>10³</th>
<th>10⁴</th>
<th>10⁵</th>
<th>10⁶</th>
<th>10⁷</th>
<th>10⁸</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda_0 = 10^{-8} )</td>
<td>21:49</td>
<td>7:10</td>
<td>2:19</td>
<td>0:44</td>
<td>0:18</td>
<td>0:11</td>
<td>0:08</td>
<td>0:07</td>
<td>0:06</td>
</tr>
<tr>
<td>( \zeta = 1 )</td>
<td>22:32</td>
<td>7:23</td>
<td>2:28</td>
<td>0:46</td>
<td>0:20</td>
<td>0:11</td>
<td>0:07</td>
<td>0:06</td>
<td>0:07</td>
</tr>
<tr>
<td>( \zeta = 50 )</td>
<td>24:51</td>
<td>8:17</td>
<td>2:39</td>
<td>0:50</td>
<td>0:20</td>
<td>0:11</td>
<td>0:07</td>
<td>0:07</td>
<td>0:06</td>
</tr>
<tr>
<td>( \zeta = 100 )</td>
<td>27:54</td>
<td>9:23</td>
<td>2:50</td>
<td>0:52</td>
<td>0:22</td>
<td>0:12</td>
<td>0:08</td>
<td>0:08</td>
<td>0:07</td>
</tr>
<tr>
<td>( \zeta = 500 )</td>
<td>52:10</td>
<td>15:20</td>
<td>4:70</td>
<td>1:30</td>
<td>0:60</td>
<td>0:35</td>
<td>0:22</td>
<td>0:21</td>
<td>0:20</td>
</tr>
</tbody>
</table>

![Fig. 9](image)

**Fig. 9.** The number of iterations needed to meet the convergence criteria during explicit integration of the systems of equations for (a) orientation and (b) reorientation as \( \zeta \) was varied.

![Fig. 10](image)

**Fig. 10.** The first quadrant of the stress responses of the Explicit Integration scheme simulations for a mass scaling factor of \( 10^8 \), plotted vs. the base (converged) solution.

4.1.3. Comparison of the schemes

Through consideration of the previously presented results (i.e., Sections 4.1.1 and 4.1.2), the most significant difference between the two schemes is that Scheme 1 must be manually calibrated and studied for convergence when mesh and/or model parameters are changed while Scheme 2 is self-calibrating, even when explicit integration of the PDEs for orientation and reorientation is turned off (i.e., \( \zeta \) set to 1). The question for a user of this constitutive model then becomes “Does the coding complexity and computational cost of calculating Eqs. (57) and (58) outweigh the benefits of self-calibration?” While the final answer to this question may be subject to both personal opinion and frequency of use of the model, in the remainder of this section we will quantify the computational cost of each scheme both with and without mass scaling for a single 3D element, and in Section 4.2 this study of numerical simplicity of Scheme 1 vs. benefits of self-calibration of Scheme 2 will be extended through simulations of a more complicated mesh.

To aid in the ensuing discussion, CPU times for the simulations both with and without mass scaling are listed in Table 4. These simulations were performed on a single core of a Dell T5500 Linux workstation. For the ensuing mass scaling study, \( \lambda_0 = 10^{-8} \) was chosen for Scheme 1 since it was in the middle of the observed convergence interval of \( 10^{-7} \leq \lambda_0 \leq 10^{-9} \) while Scheme 2 was studied for both no integration, i.e., \( \zeta = 1 \), and \( \zeta = 50 \). Mass scaling was imposed by use of a mass scaling factor \( [8] \), which was varied in order of magnitude from 1 (no mass scaling) to \( 10^8 \).

Comparing the results for no mass scaling, Scheme 2 without integration (\( \zeta = 1 \)) took 3% longer to run. Both schemes provided converged stress, strain, and volume fraction calculations, but neither scheme provided a converged Lagrange multiplier solution (see the preceding two sections and Figs. 5–8). However, the Scheme 2 Lagrange multiplier calculations for \( \zeta = 1 \) were of the same order of magnitude as the converged solution (Fig. 8), while the Scheme 1 fixed Lagrange multiplier increment was off by an order of magnitude. Integration of the PDEs using \( \sim 50 \) iterations increased the time by an additional 6%, while providing the additional accuracy of a converged Lagrange multiplier solution. Further increase of the number of iterations for each integration continued to increase the computational cost with virtually no benefit, as the Lagrange multiplier calculations did not experience substantial improvement (Fig. 8c).

As mentioned in the introduction to Section 4, since this constitutive model is rate independent, it is desirable to have a numerical implementation that is insensitive to mass scaling so this technique may be invoked to reduce the computational cost of simulations. Stress responses of Scheme 1 as the mass scaling factor is varied

![Table 4](image)

**Table 4**

<table>
<thead>
<tr>
<th>Factor</th>
<th>None</th>
<th>10</th>
<th>10²</th>
<th>10³</th>
<th>10⁴</th>
<th>10⁵</th>
<th>10⁶</th>
<th>10⁷</th>
<th>10⁸</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda_0 = 10^{-8} )</td>
<td>21:49</td>
<td>7:10</td>
<td>2:19</td>
<td>0:44</td>
<td>0:18</td>
<td>0:11</td>
<td>0:08</td>
<td>0:07</td>
<td>0:06</td>
</tr>
<tr>
<td>( \zeta = 1 )</td>
<td>22:32</td>
<td>7:23</td>
<td>2:28</td>
<td>0:46</td>
<td>0:20</td>
<td>0:11</td>
<td>0:07</td>
<td>0:06</td>
<td>0:07</td>
</tr>
<tr>
<td>( \zeta = 50 )</td>
<td>24:51</td>
<td>8:17</td>
<td>2:39</td>
<td>0:50</td>
<td>0:20</td>
<td>0:11</td>
<td>0:07</td>
<td>0:07</td>
<td>0:06</td>
</tr>
<tr>
<td>( \zeta = 100 )</td>
<td>27:54</td>
<td>9:23</td>
<td>2:50</td>
<td>0:52</td>
<td>0:22</td>
<td>0:12</td>
<td>0:08</td>
<td>0:08</td>
<td>0:07</td>
</tr>
<tr>
<td>( \zeta = 500 )</td>
<td>52:10</td>
<td>15:20</td>
<td>4:70</td>
<td>1:30</td>
<td>0:60</td>
<td>0:35</td>
<td>0:22</td>
<td>0:21</td>
<td>0:20</td>
</tr>
</tbody>
</table>

![Fig. 9](image)

**Fig. 9.** The number of iterations needed to meet the convergence criteria during explicit integration of the systems of equations for (a) orientation and (b) reorientation as \( \zeta \) was varied.

![Fig. 10](image)

**Fig. 10.** The first quadrant of the stress responses of the Explicit Integration scheme simulations for a mass scaling factor of \( 10^8 \), plotted vs. the base (converged) solution.
Fig. 11. Total kinetic energy of single 3D element simulations as mass scaling factor was varied: (a) Fixed Lagrange Multiplier scheme with $\lambda_0 = 10^{-3}$ (see Section 3.1); (b) Explicit Integration scheme with $\zeta = 1$ (see Section 3.2); (c) Explicit Integration scheme with $\zeta = 50$ (see Section 3.2).

Fig. 12. Contours of the evolution of the oriented martensite volume fraction as the specimen is loaded.
are plotted in Fig. 7c and d. Mass scaling for this scheme was robust for factors less than or equal to $10^3$. Contrarily, stress, strain, and volume fraction were correctly calculated using Scheme 2 for all mass scaling factors except for $10^5$. For $10^5$, the predictions were still quite reasonable on the whole, though locally there were some oscillations near load inflection points. The imparting of dynamic responses upon the quasi-static solution indicates that the threshold for increasing the mass scaling factor without affecting the rate independent calculations had been crossed (see Fig. 10). Fig. 10 also shows that these oscillations were more quickly dampened for full integration ($\zeta = 50$) than for self-calibration only ($\zeta = 1$).

The total kinetic energies of each simulation of the mass scaling study are plotted in Fig. 11. Note that for a mass scaling factor of $10^3$, this plot would suggest both schemes performed equivalently. However, recall that Scheme 1 calculations converged only for factors less than or equal to $10^3$, and in Fig. 11 Scheme 1 kinetic energy oscillates more severely than Scheme 2 for this range of mass scaling factors. Furthermore, note that while the average kinetic energy of self-calibrating Scheme 2 does increase for increasing mass scaling factor, it is relatively oscillation-free for factors of $1–10^7$, all of which provided converged calculations of stress, strain, and volume fraction. Clearly the self-calibrating scheme is more robust with regard to retaining rate independent, quasi-static conditions in the presence of mass scaling.

Concurrent consideration of all of these results demonstrates that the additional computational expense of making the self-calibration calculation for the initial Lagrange multiplier increments using Eqs. (57) and (58) is quickly recovered in the robustness of Scheme 2 in the presence of mass scaling. Converged simulations using Scheme 2 were obtained in 6–7 s with a mass scaling factor of $10^5$ while the least expensive converged simulation for Scheme 1 with $\lambda_0 = 10^{-8}$ took 44 s (6–7 times longer) with a mass scaling factor of $10^3$.

### 4.2. Simulations of a 2D mesh with varying element size

To evaluate the computational efficiency of the schemes in a multi-element simulation, a 3 mm x 25 mm bar with a 0.5 mm diameter hole placed at the center of the gage area was subjected to an axial displacement control load of 0.8 mm at a rate of 0.1 mm/s. The model was seeded with 28 equidistant nodes around the diameter of the hole, the seeding was biased along the y-axis, top, and bottom edges such that a non-uniform mesh of CPS4R [8] 4-node plane-stress elements that was coarse away from the hole and fine near the hole was achieved. This mesh is shown in Figs. 12 and C.1.

The CPU times and computational parameters ($\lambda_0$, $\zeta$, mass scaling factor) of the simulations used for this study are listed in Table 5. The same $\lambda_0$ and $\zeta$ values used in the mass scaling study of the previous section were selected, and the mass scaling factor was chosen to be one order of magnitude less than the maximum mass scaling factor that resulted in a converged solution for the single 3D brick element (see Section 4.1.3). Immediately the benefit of making the extra calculation to provide self-calibration is evidenced in noting that the Scheme 1 simulation was on the order of 115 h while the Scheme 2 simulations completed in ~4.5 h. Furthermore, the additional computation of integrating the PDEs in Steps 4 & 5 using ~50 explicit iterations does not substantially alter the CPU time of the simulation.

Similar to the single element studies, choices of $\zeta = 1$ and $\zeta = 50$ produced fully converged solutions for stress, strain, and volume

---

**Table 5**

<table>
<thead>
<tr>
<th>Description</th>
<th>CPU Time (h:m:s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_0 = 10^{-8}$, Factor = $10^3$</td>
<td>115:07:24</td>
</tr>
<tr>
<td>$\zeta = 1$, Factor = $10^2$</td>
<td>4:35:41</td>
</tr>
<tr>
<td>$\zeta = 50$, Factor = $10^3$</td>
<td>4:41:34</td>
</tr>
</tbody>
</table>

---

**Fig. C.1.** (a) Isolated responses of each loading segment in equivalent stress–strain space are shown along with (b) oriented martensite volume fraction and (c) orientation and reorientation Lagrange multiplier evolutions over simulation time. This figure complements the Appendix C discussion of SMA mechanics during the converged biaxial superelastic loading responses of Section 4.1.
fraction calculations. Snapshots of oriented martensite volume fraction contours are shown in Fig. 12. The simulation using Scheme 1 also converged to these solutions.

5. Conclusions

The numerically improved constitutive model demonstrated robustness for a complex, non-proportional loading in a single 3D brick element, and also in predicting transformation about a geometric defect (i.e., hole) within a rectangular gage section. Though the focus of this work was to evaluate the best method for numerical implementation of this constitutive model within Abaqus Explicit software [8], it may be noted that the phase transformation propagation patterns shown in Fig. 12 are qualitatively consistent with empirical observations of other researchers [35]. The transformation nucleates in lobe-like formations near the defect, propagates away from the defect in a crossing pattern, and in the bulk material bands form along shear planes and meld into regions of uniformly transformed material. This indicates that the improved model is not only numerically robust, but also able to realistically replicate SMA behaviors.

Two schemes for solving the PDEs of the model were derived. In Scheme 1, constant Lagrange multiplier increments were used and the error in the VUMAT calculation was compensated by the Abaqus Explicit FE calculation. In Scheme 2, a newly formulated explicit integration scheme was used. While this scheme requires extra computations in the VUMAT, it demonstrated a superior numerical robustness relative to Scheme 1, highlighted by the following findings:

1. Scheme 2 is self-calibrating, thus a convergence study of the VUMAT for the choice of ε0 using a guess-and-check method is not required for Scheme 2 as it is for Scheme 1.
2. Solutions using Scheme 2 converged over 7 orders of magnitude of mass scaling factor, while Scheme 1 calculations only converged for 3.
3. This robustness in sensitivity to mass scaling factor results in Scheme 2 being the less expensive of the two schemes in multi-element simulations, demonstrated by a simulation time of 4.5 h vs. 115 h for Scheme 1 in calculating the stress-induced transformation response of a bar with a hole.

Furthermore, in studying the choice of ζ = 1 for Scheme 2 it was found that the stress, strain, and martensite volume fraction predictions converged equally as well as choice of ζ = 50 in all of the simulations. The only benefit to allowing for multiple iterations in the integration is if accurate Lagrange multiplier increment values are desired. However, since in this model they are merely a numerical tool for integrating PDEs and have no physical relevance, there is rarely a need to calculate them exactly as the Abaqus Explicit FE solver will compensate for error as long as they are estimated within reason (see Section 4.1.2).

If we now consider these two schemes relative to the original, unaugmented scheme proposed by Panico and Brinson in which Newton’s method was used to implicitly integrate the PDEs, Scheme 2 with ζ = 1 requires no iteration to achieve a converged solution for stress, strain, and volume fraction calculations within an explicit FE framework. Thus, the true essence of an explicit FE simulation is recovered with Scheme 2 – in solving the material model, make a single calculation that is close but not necessarily exact for each time step and let the explicit FE calculation compensate for the error. Even though we did not create a VUMAT using an unaugmented version of the originally proposed scheme, clearly Scheme 2 with ζ = 1 is less computationally expensive than implicitly integrating the equations at each step of and explicit FE simulation. Additionally, while Eqs. (57) and (58) are not simple, they are much easier to program than the full Jacobian required for the implicit integration scheme (see Appendix B). Hence, if implementing this constitutive model into explicit finite elements for studies using mass scaling, Scheme 2 with ζ = 1 is the best choice, and Scheme 2 with ζ = 50 may be used to ensure a convergent solution in all variables.

In considering impacts of this work beyond our model, a similar derivation approach as was used in Section 3.2 may be used to derive an integration-free, self-calibrating explicit scheme for other material constitutive models: simply formulate the relationship between driving forces and kinetic equations through Kuhn-Tucker conditions, then derive the initial Lagrange multiplier increment following our procedure setting ζ = 1. Finally, we note that many other SMA FE modelers chose to use implicit FE solvers (e.g., [42,29,36,37]), such as Abaqus Standard [8]. In an implicit finite element framework where implicit integration is an inherent part of the solution scheme, the Jacobian in Appendix B will be useful in formulating the tangent matrix [8]. While our long-term vision for developing an SMA constitutive model using this core model led us to select an explicit FE implementation for model development, a purely implicit FE implementation may be equally robust as the explicit Scheme 2 presented in this article for this rate and history independent improved model.

Acknowledgements

A.S. acknowledges funding through fellowships from the Toshio Mura Endowment, Predictive Science and Engineering Design

Table A1
Superscripts and subscripts.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AM</td>
<td>austenite, martensite</td>
</tr>
<tr>
<td>e, tr</td>
<td>elastic, transformation</td>
</tr>
<tr>
<td>σ, r, SA</td>
<td>orientation, reorientation, self-accommodation</td>
</tr>
<tr>
<td>f</td>
<td>forward, reverse</td>
</tr>
<tr>
<td>s, f</td>
<td>start, finish</td>
</tr>
<tr>
<td>0</td>
<td>indicates a stress-free, reference, or initial condition, as defined in text</td>
</tr>
<tr>
<td>n</td>
<td>denotes value from previous step</td>
</tr>
<tr>
<td>f = 1</td>
<td>if not used in time-discrete equations, the represented quantity is in the n + 1 (current) condition</td>
</tr>
</tbody>
</table>

Table A2
Variables.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ</td>
<td>Stress</td>
</tr>
<tr>
<td>σ'</td>
<td>Deviatoric stress</td>
</tr>
<tr>
<td>σ̄</td>
<td>Mises equivalent stress</td>
</tr>
<tr>
<td>ε</td>
<td>Strain</td>
</tr>
<tr>
<td>ψ</td>
<td>Helmholtz free energy</td>
</tr>
<tr>
<td>T</td>
<td>Temperature</td>
</tr>
<tr>
<td>ζ</td>
<td>Martensite volume fraction</td>
</tr>
<tr>
<td>u</td>
<td>Specific energy</td>
</tr>
<tr>
<td>η</td>
<td>Specific entropy</td>
</tr>
<tr>
<td>i</td>
<td>Lagrange multiplier</td>
</tr>
<tr>
<td>Dp</td>
<td>Dissipation</td>
</tr>
<tr>
<td>Y</td>
<td>Dissipative/frictional force</td>
</tr>
<tr>
<td>X</td>
<td>Thermodynamic driving force</td>
</tr>
<tr>
<td>F</td>
<td>Limit function</td>
</tr>
<tr>
<td>̃ι</td>
<td>Projection tensor (See [41])</td>
</tr>
</tbody>
</table>
Cluster at Northwestern (PSED), Initiative for Sustainability and Energy at Northwestern (ISEN). A.S. and C.B. acknowledge the support of the Army Research Office, Grant # W9111NF-12-1-0013/PO0002. Catherine Tupper is thanked for creating the model of the beam with the hole as well as many fruitful discussions and testing development versions of the VUMAT. Pingping Zhu is thanked for continuing validation efforts for the VUMAT.

Appendix A

The following conventions are used in this work:

Scalar—not bold font

2nd Order Tensor—bold font

4th Order Tensor—capital open-face font

If $A$ is a tensor then: $AA = A_{ij}A_{ik}$, $A : A = A_{ij}A_{ij}$,

$A^t = A_{ij} - \frac{1}{2} \delta_{ij} A_{kk}$

The following shorthand is used:

$N(A) = \frac{A}{\|A\|}$

$|\arg| = \frac{\arg + |\arg|}{2}$

$|\arg|$ is the Euclidean norm (Tables A.1–A.4).

Appendix B

It was originally proposed that Newton–Raphson [38] be used to implicitly integrate the orientation and reorientation equations [1]. This scheme achieves quadratic convergence, thus less iterations are needed to solve the coupled equations than in the full Explicit Integration scheme described in this paper. However, the need to program the Jacobian equations, presented in this appendix, makes it more tedious to implement.

The algorithm as applied to the improved model is presented followed by the systems of unknowns $\{x\}$, equations $\{a\}$, and Jacobians $\{D\}$ that must be solved using this algorithm; $\{\Delta x\}$ is the increment of the unknowns calculated during each iteration.

\[
\{x_0\} \text{ (initial guesses) = values calculated in the elastic trial step, or from the previous step if they were not changed in Steps 1–4.}
#iterations = 1
While residual of $\{a\}$ is less than a convergence tolerance and #iterations ≤ max iterations
\begin{itemize}
  \item Evaluate $\{a\}$ and $\{D\}$ using $\{x_{current}\}$
  \item Solve $\{\Delta x\} = -\{D\}^{-1} : \{a\}$.
  \item Increment $\{x_{current}\} = \{x_{current}\} + \{\Delta x\}$.
  \item Calculate residual $\{a\}$.
  \item #iterations = #iterations + 1
\end{itemize}
End

The presentation of the equations to solve ensues, incorporating the additional shorthand:

\begin{equation}
\frac{1}{\sqrt{2 \epsilon_{ir}^{max}}} h(T, \xi_\sigma) = \frac{\rho((T_{\text{fin}} - \tilde{\epsilon}) + H_n \tilde{\epsilon}_{\xi_\sigma})}{\sqrt{2 \epsilon_{ir}^{max}}} \tag{B.1}
\end{equation}

\begin{equation}
\frac{dY_{\xi_\sigma}}{d\xi_{\sigma}} = \left\{ \begin{array}{ll}
\tan(\pi(\xi_\sigma + \frac{1}{2})) + \frac{\pi}{\cos(\pi(\xi_\sigma + \frac{1}{2}))}, & \xi_\sigma > 0 \\
(\frac{2}{\pi^2})^{\frac{1}{2}} (\frac{\pi}{\sin(\pi(\xi_\sigma + \frac{1}{2}))})^{\frac{1}{2}} + \frac{\pi}{\cos(\pi(\xi_\sigma + \frac{1}{2}))}, & \xi_\sigma > 0
\end{array} \right. \tag{B.2}
\end{equation}

B.1. Scenario i: Orientation

The number of equations to solve may be reduced from the 20 presented in Eqs. (23)–(27) to the 14 stated here by substituting Eqs. (23)–(27) into Eq. (24) and eliminating $X_{B}:
\{x\} = \{\sigma, \xi_{\sigma}, \Delta \lambda_{\sigma}, \xi_{\sigma}\}
\{a\} = \left\{ \begin{array}{c}
\sigma - C : (\varepsilon - \varepsilon_{ir}) \\
\varepsilon_{ir} - \Delta \lambda \sigma \varepsilon_{ir} - h(T, \xi_{\sigma}) N(\sigma) \\
F_{a} - ||\sigma|| + h(T, \xi_{\sigma}) + Y_{\sigma}(\xi_{\sigma}) \\
\varepsilon_{\xi_{\sigma}} - \frac{|\varepsilon_{\xi_{\sigma}}|}{\sqrt{2 \epsilon_{ir}^{max}}}
\end{array} \right. \quad \{D\} = \left[ \begin{array}{cccc}
1 & C & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{\rho(T_{\text{fin}} - \tilde{\epsilon}_{\sigma}) + H_n \tilde{\epsilon}_{\xi_{\sigma}}}{\sqrt{2 \epsilon_{ir}^{max}}} \\
0 & 0 & 0 & 1
\end{array} \right]
\end{equation}

B.2. Scenario ii: Reorientation

In solving the reorientation equations, the desired outputs are $\sigma_{ref}$. Eqs. (12) and (13) may be consolidated, reducing the number of equations to integrate to the 13 shown below. $\Delta \lambda_{\sigma}$ may then be calculated post-integration if explicit tracking of its evolution is desired.

$\{x\} = \{\sigma, \xi_{\sigma}, \Delta \lambda_{\sigma}\}$

Table A.3

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E$</td>
<td>Effective Young's modulus</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Effective Poisson's ratio</td>
</tr>
<tr>
<td>$C$</td>
<td>Isotropic elasticity tensor</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Density</td>
</tr>
<tr>
<td>$\tilde{\epsilon}_{\text{ir}}$</td>
<td>Specific entropy of phase transformation $\tilde{\epsilon}_{\text{ir}} - \tilde{\epsilon}^{\text{ref}}$ (Equation of state was used in [1])</td>
</tr>
<tr>
<td>$\tilde{\epsilon}$</td>
<td>Specific energy of phase transformation $\tilde{\epsilon} - \tilde{\epsilon}_{\text{ir}}$ (Equation of state was used in [1])</td>
</tr>
<tr>
<td>$C_{p}$</td>
<td>Specific heat at constant volume</td>
</tr>
<tr>
<td>$M_{T}$</td>
<td>Martensite start temperature</td>
</tr>
<tr>
<td>$M_{F}$</td>
<td>Martensite finish temperature</td>
</tr>
<tr>
<td>$A_{S}$</td>
<td>Austenite start temperature</td>
</tr>
<tr>
<td>$A_{F}$</td>
<td>Austenite finish temperature</td>
</tr>
<tr>
<td>$C_{AM}$</td>
<td>Paté–Cohen coefficient</td>
</tr>
<tr>
<td>$Y_{\text{max}}$</td>
<td>Dissipative(frictional) constant</td>
</tr>
<tr>
<td>$\omega_{e}$</td>
<td>Maximum effective transformation strain $\gamma$ [1]</td>
</tr>
<tr>
<td>$H_{\sigma}$</td>
<td>Hardening coefficient for orientation</td>
</tr>
<tr>
<td>$A_{\text{ref}}$</td>
<td>Austenite parameters</td>
</tr>
<tr>
<td>$C_{\text{ref}}$</td>
<td>Kinetic parameters $\theta$ [1]</td>
</tr>
<tr>
<td>$k^f$</td>
<td>A third kinetic parameter $B$ was also used in [1]</td>
</tr>
<tr>
<td>$C_{\text{ref}}$, $Y_{\gamma_{0}}$</td>
<td>Internal parameters that relate transformation temperatures with specific energy and entropy</td>
</tr>
</tbody>
</table>

Table A.4

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tilde{\epsilon}_{\text{ir}}$</td>
<td>Onset strain sensitivity parameter</td>
</tr>
<tr>
<td>$\lambda_{\sigma}$</td>
<td>Lagrange multiplier increment for Scheme 1</td>
</tr>
<tr>
<td>$\lambda_{\sigma}$</td>
<td>Targeted number of iterations for Scheme 2</td>
</tr>
</tbody>
</table>

End
the stress-space response shown in Figs. 5b and 6b, some of which this figure will provide insight into the inflections labeled i–v of increments for each orientation and reorientation. Our review of loading segment documented in Table 3; (b) the oriented martensite-equivalent stress–strain response to each displacement-controlled

4.1. In this discussion, we refer to Fig. C1, which shows (a) the oriented martensite or both mechanisms are active. The oriented martensite does not correlate with inflections in the loading path.

Observing each orientation and reorientation Lagrange multipliers allow the non-proportional addition of shear to the axial load state to further increment the material for forward through phase transformation in equivalent stress–strain space (Fig. C1a).

Path ii to iii: Very quickly into this step of continued axial unloading through zero, reverse transformation of oriented martensite to austenite begins and reorientation turns off (Fig. C1b and c). Thus the material begins to proceed through the reverse-transformation portion of a minor hysteresis loop (i.e. a lower plateau) in equivalent stress–strain space as calculations are made to satisfy Eqs. (26)–(29) (Fig. C1a; see [18,39,40] for further discussion of minor transformation hysteresis loops in SMAs). Simultaneous to reverse transformation without reorientation, $\hat{\mathbf{v}}(\mathbf{e}_r): \mathbf{\sigma}$ grows since the non-proportional loading is driving the deviatoric stress in a new direction. Eventually, this quantity grows to $\sigma_{re} = 40$ MPa and through Eqs. (19) and (22) reorientation turns on again and redirects the transformation strain until it is aligned with the deviatoric stress. However, reorientation turning on and off alone does not cause sharp inflections in the stress response, but instead creates a nonlinear “rounding” effect when coupled with the elastic response in the equivalent stress and strain as the tensile load diminishes and the axial loading becomes compressive.

Path iii to iv: The inflection iii observed in Fig. 6b, which results in forward evolution of oriented martensite (Fig. C1b). Only orientation is active during this segment (Fig. C1c).

10 to 20: Fixing axial displacement, shear displacement of 1.6 mm is applied to the nodes of the top surface of the brick element (Table 3). This triggers reorientation (Fig. C1c) since the deviatoric stress state changes directions, but notice there is a slight lag in the reorientation response. This is because $\sigma_{re}$ is 40 MPa, not 0. Thus dissipation must occur through Eq. (19) before reorientation activates. Here this dissipation occurs very quickly due to the sharp change in the applied load direction. Orientation turns off for a very short time then slowly increases (Fig. C1c). Thus, further orientation of martensite occurs, but it is slow and small in magnitude (Fig. C1b). These mechanisms allow the non-proportional addition of shear to the axial load state to further increment the material forward through phase transformation in equivalent stress–strain space (Fig. C1a).

20 to 40: Path i to ii: Axial displacement of –3.2 mm is applied to the nodes of the top surface of the brick element (Table 3). This action initially halts evolution of oriented martensite, which was active at the end of the previous step; however, reorientation of inelastic strain continues (Fig. C1c). There is a large drop in equivalent stress (Figs. 5b, 6c, Fig. C1a) as elastic strain is relaxed while the magnitude of inelastic strain remains constant (again, through Eqs. (6) and (7) reorientation does not affect the magnitude of transformation strain). However, because the transformation strain is changing direction, the slope of the equivalent stress–strain response i–ii is much steeper than that of the initial elastic loading in time segment 0 to 10 (Figs. 6c and C1a).

Appendix C

Here we provide a more detailed review of the mechanics at play in the converged response of the non-proportional strain control loading of Fig. 6 imposed upon the brick element in Section 4.1. In this discussion, we refer to Fig. C1, which shows (a) the equivalent stress–strain response to each displacement-controlled loading segment documented in Table 3; (b) the oriented martensite volume fraction evolution in time; (c) the Lagrange multiplier increments for each orientation and reorientation. Our review of this figure will provide insight into the inflections labeled i–v of the stress-space response shown in Figs. 5b and 6b, some of which do not correlate with inflections in the loading path.

Observing each orientation and reorientation Lagrange multiplier increments allows us to examine when orientation, reorientation, or both mechanisms are active. The oriented martensite volume fraction provides understanding as to whether oriented martensite is evolving or de-evolving, as this volume fraction is completely independent of reorientation (see Section 1.2). When concurrently considered with these responses, the equivalent stress–strain response will allow us to understand axial-shear couplings that may not be obvious in examining the response in axial-shear space (i.e. Fig. 6). We now proceed to explain the mechanics of each loading segment.

0 to 10: An axial displacement of 1.6 mm is applied to the top surface of the $40 \times 40 \times 40$ mm$^3$ brick element while the bottom surface is fixed (Table 3). Initially, the loading is modeled as elastic with a 70 GPa effective Young’s modulus. It reaches a critical stress of ~325 MPa and beings forward phase transformation (Fig. C1a), which results in forward evolution of oriented martensite (Fig. C1b). Only orientation is active during this segment (Fig. C1c).
of forward transformation in the minor loop (Fig. C1). In between, reorientation is the only active inelastic deformation mechanism, and similar to path i-ii, it only redirects the inelastic strain but does not alter its magnitude (Section 1.2, Eq. (7)) and the slope of the equivalent stress and strain loading again deviates from the elastic loading response as a result, exhibiting a much harder effective modulus. Path iv to i: Finally, the forward transformation proceeds along the global hysteresis upper transformation plateau along path iv–i until the –3.2 mm axial displacement is complete. Note that the plateau is initially overshoot a bit at point iv, Fig. 6c due to the dissipation Y_s that must occur before forward transformation re-initiates (Eq. (20)). Reorientation is initially active and decreasing during this path, and it is through this mechanism that the global equivalent stress–strain response is redirected after the overshoot (Fig. C1c). Once the transformation strain and deviatoric stress directions are realigned, reorientation ceases (Fig. C1c) and this path concludes in a forward-transformation to oriented martensite manner (Fig. 6c and Fig. C1).

40 to 60: Shear displacement of –3.2 mm is applied to the nodes of the top surface of the brick element (Table 3). This action immediately triggers reorientation and turns off orientation (Fig. C1c), thus the oriented martensite volume fraction does not evolve during this load path (Fig. C1b). Contrary to the previous load segment, the effective modulus of combined elastic and martensite reorientation segment of the global hysteresis loop in equivalent stress–strain space is more compliant than the elastic-only response (Fig. 6c and Fig. C1a). The reason it is stiffer in Path i–ii and more compliant in Path i–v is because of the nature of the reorientation and the calculation of equivalent strain. Observe from Eq. (60) that in isolation, reorienting the transformation strain from an axial alignment to a shear alignment will increase the effective strain, while the reverse action will decrease the effective strain as the shear component is weighted by 4/3 relative to the unity weighting of the axial component in this calculation. With this understanding, considering the changes in the loading path (Fig. 6a): while the elastic stress–strain response is always constant in this model, in Segment 20 to 40 the reorientation event as axial strain diminished while shear displacement was held constant pointed the transformation strain from an axial-biased toward a more shear-biased direction. Even though the magnitude of the inelastic strain was constant the calculation of effective inelastic strain increased through Eq. (60), which combined with the elastic response caused the material to appear effectively stiffer. In this segment (40–60), however, the opposite effect is observed since reorientation is triggered by a change in shear strain while axial displacement is held constant, biasing the inelastic shear direction from a shear alignment toward an axial alignment, which decreases the calculation of effective transformation strain through Eq. (60).

The material requirement of following a constant volume fraction response is why axial stress decreases along with shear stress in Fig. 6b. The inflection v observed in Fig. 6b and c, Fig. C1a during this loading segment is a result of crossing the 0 mm shear displacement in unloading to reloading. Reloading occurs along the same constant volume fraction response in equivalent stress–strain space as the shear is increased to –1.6 mm displacement in the shear direction (Fig. C1a). Again, the reason the inflection point occurs at σ_1 = –40 MPa instead of 0 MPa is because of the dissipation necessary to reorient martensite (Eq. (19)).

60 to 80: Axial displacement of 3.2 mm is applied to the nodes of the top surface of the brick element (Table 3), returning this surface to global 1.6 mm axial displacement. The mechanisms at play during this load segment are identical as from 20 to 40. Their effective response in equivalent stress–strain space is similar but not identical, again due to dissipative terms in Eqs. (18)–(20) that provide phenomenological energy barriers that must be overcome for transformation, orientation, and reorientation of martensite.

80 to 90: Shear displacement of 1.6 mm is applied to the nodes of the top surface of the brick element, returning the surface to 0 mm displacement in the shear direction (Table 3). Thus, again reorientation of a constant volume fraction of martensite and elastic unloading concurrently relax the material at a slope that is more compliant than the elastic-only response (path i–v, Fig. C1a, Fig. 6c). Again, in Fig. 6b the asymmetry about σ_1 = 0 MPa in this response is due to energy dissipation.

90 to 100: Finally axial displacement of –1.6 mm is applied to the nodes of the top surface of the brick element, returning the surface to 0 mm displacement in all directions (Table 3). Through relatively quick energy dissipation processes, the shear stress is relaxed (Fig. 6b and c, Fig. C1a). The remainder of this displacement allows the material to complete reverse phase transformation and finally elastic unloading in a uniaxial stress state (Fig. C1a, Fig. 6b and c). This is done through the orientation mechanism of the model (Fig. C1b and c). The reason there is a smooth transition from transformation to elasticity in the reverse transformation as opposed to the sharp transition that was observed in the forward transformation is because of the form of Eq. (20); i.e., this smoothing is merely a product of the tangent function used to impose the bounds upon the volume fraction.

References
