MIT Open Access Articles

Multiphysics design of programmable shape-memory alloy-based smart structures via topology optimization

Massachusetts Institute of Technology

Multiphysics design of programmable shape-memory alloy-based smart structures via topology optimization

Multiphysics Design of Programmable Shape-Memory Alloy-Based Smart Structures via Topology Optimization

Ziliang Kang, Kai A. James

University of Illinois at Urbana-Champaign, Department of Aerospace Engineering

Urbana, Illinois, United States

Abstract

We present a novel multiphysics and multimaterial computational design framework for shape-memory alloy-based smart structures. The proposed framework uses topology optimization to optimally distribute multiple material candidates within the design domain, and leverages a non-linear phenomenological constitutive model for shape-memory alloys (SMAs), along with a coupled transient heat conduction model. In most practical scenarios, SMAs are activated by a nonuniform temperature field or a nonuniform stress field. This framework accurately captures the coupling between the phase transformation process and the evolution of the local temperature field. Thus, the resulting design framework is able to optimally tailor the two-way shape memory effect and the superelasticity response of SMAs more precisely than previous algorithms that have relied on the assumption of a uniform temperature distribution. We present several case studies, including the design of a self-actuated bending beam and a gripper mechanism. The results show that the proposed framework can successfully produce SMA-based designs that exhibit targeted displacement trajectories and output forces. In addition, we present an example in which we enforce material-specific thermal constraints in a multimaterial design to enhance its thermal performance. In conclusion, the proposed framework provides a systematic computational approach to consider the nonlinear thermomechanical response of SMAs, thereby providing enhanced programmability of the SMA-based structure.

Keywords: shape-memory alloys, multimaterial design, transient heat conduction, two-way shape memory effects, superelasticity, topology optimization, programmable smart structures

1. Introduction

As the most widely adopted thermoelectrical active materials, shape-memory alloys (SMAs) have been the subject of extensive research for decades. SMAs are capable of reversible deformation under a temperature cycle or a loading cycle, referred as the two-way shape memory effect (TWSME) and superelasticity (or psuedoelasticity) respectively. These two phenomena occur as a result of the diffusionless phase transformation between the martensite phase and austenite phases. As a type of metallic material, SMAs have high-pressure and high-temperature resistance [1, 2]. These characteristics distinguish SMAs from most active materials which are polymeric, therefore SMA can be used in a broad range of applications, including aerospace, automotive, robotics applications, as well as biomedical devices and integrated rescue equipment [1, 3]. In addition, SMAs can be actuated by electric stimuli through Joule heating, enabling more choices of activation approaches for SMA-based smart structures [3].

For years, the design techniques used for SMA-based structures relied on a combination of experiments and simulation, in which the shape and topology of the design was obtained heuristically

[3, 4, 5]. On the simulation side, finite element analysis is the main tool used in modeling the material response of SMAs. Researchers have been developing multiple constitutive models for SMAs in recent years, mainly categorized into micromechanical models and phenomenological models. Sun and Hwang developed early micromechanical models [6, 7], which were later built upon by Bhattacharya [8, 9]. These models are preferred in describing the micromechanics behaviors of SMAs varied with different lattice orientation. Further investigation considering the differences between lattice structures of SMAs, including twining, detwining, and single crystallization [10, 11], was conducted as well. Meanwhile, the history of developing phenomenological models has focused on proposing proper thermodynamic hardening models to capture the latent heat exchange process of phase transformation. The most popular hardening models include the exponential model [12], the cosine model [13, 14], the quadratic model [5, 15, 16] and the smooth transformation model [17], On the experimental side, extensive research has been conducted to enhance the performance of SMAs by increasing their bandwidth [18, 19], fatigue life [20, 21], and stability [3]. Systematic efforts have also been directed toward enhancing the attributes of SMAs to allow for a wider range of operating temperatures by introducing new material compositions, such as Nickel-rich NiTi SMAs, Ti-ta alloys, and Ni-Mn-Ga alloys [22, 23, 24, 25].

Despite ongoing efforts, limitations of SMAs still exist, including relatively small usable strain, low actuation frequency, low controllability, and low energy efficiency [3]. The major obstacle, low operational frequency and narrow bandwidth of SMAs, are results of inefficient heat transfer into and out of the active materials [1, 3]. While research shows that Joule heating with high electrical currents can increase the heating rate, SMA-based structures still faces problems of overheating and damage to the actuator [3, 26, 27]. In addition, low heat conduction efficiency is the most significant factor limiting bandwidth during the cooling process [3, 28, 29]. With the classical experiment-simulation design approaches restricting the shape of SMA-based smart structure to rods, plates, ribbons, springs and wires [1], the lack of geometric design freedom limits the performance of SMA-based actuators in terms of response time. Researchers have also proposed to enhance the thermomechanical attributes of SMAs by introducing multimaterial structures [1]. However, it again requires months of time to conduct experiments and simulations to arrive at a conceptual design. In response to those challenges, researchers seek to leverage computational design techniques to shorten the design cycle, and leverage the multi-physics properties of SMAs as much as possible. The earliest research on topology optimization of thermally-actuated compliant mechanisms was conducted by Sigmund [30, 31], which considered thermal expansion and multimaterial design. Later, Li et al. studied topology optimization of compliant mechanisms considering pure thermal expansion influenced by transient heat transfer [32]. Cho et al. proposed an efficient topology optimization scheme for thermoelastic problems using coupled field adjoint sensitivity analysis [33]. Studies in this area were also reported on topology optimization of other thermally-responsive smart materials including liquid crystal elastomers [34] and shape memory polymers [35, 36]. For SMAs, Langelaar proposed a topology optimization algorithm considering a piecewise linear constitutive material model [37]. In addition, the authors' previous work studied topology optimization of single-material SMA structures, using an inelastic model that accounts for the phase transformation of SMAs and assumes uniform temperature distribution [38]. However, multi-material design of SMA-based smart structures that enables complex motion with less active materials, as well as nonlinear thermal-responsive behaviors of smart materials coupled with transient heat conduction have yet to be rigorously explored.

In this paper, we propose a systematic computational design framework for multimaterial design of SMA-based smart structures, via the technique of topology optimization. We consider an accurate phenomenological model for SMAs [5], coupled with transient heat conduction. No assumptions of the topology or shape of the smart structures are made before the implementation of

2

topology optimization. On the contrary, the topology of the structure is determined by distributing material candidates accordingly to obtain optimal multiphysical responses of the structure, creating a design that can be manufactured through 3D printing. To the best of our knowledge, this is the first topology optimization framework for SMA-based structures that allows for multimaterial design. Further, this is also the first topology optimization framework for thermally-responsive smart materials that considers transient heat conduction. The novel features of the framework can be summarized as follows: 1) the proposed algorithm captures the interaction between the inelastic behaviors of SMAs and transient thermal conduction 2) consideration of evolution of transformation temperature and stress in terms of various environmental conditions are embedded in the topology optimization algorithm 3) a novel SIMP-based scheme is used for interpolating the thermal and mechanical properties of material candidates, which enables an efficient multimaterial design, and 4) the proposed algorithm is able to consider material-specific thermal constraints to enhance the thermomechanical attributes of SMA-based smart structures, to address the major challenge of heat conduction problems in designing SMA-based structures. In short, the proposed algorithm provides an efficient design pathway to utilize the inelastic behaviours of SMAs, while simultaneously accounting for the multimaterial composition and heat conduction performance of the structures. The proposed framework is expected to be a promising tool to enhance the structural programmability and thermomechanical attributes of SMA-based structures.

2. Methodology

2.1. Phenomenological Constitutive Relationship of SMAs

In this paper, we use the phenomenological constitutive model proposed by Boyd and Lagoudas [15] and later described by Lagoudas in 2008 [5] to define the diffusionless phase transformation of SMAs from martensite (M) to austenite (A). Note that this model assumes small deformations. Lagoudas uses the Gibbs-free energy in Equation 1 to define the thermodynamic potential of SMAs, since it contains the essential internal state variables of stress tensor $\boldsymbol{\sigma}$ and temperature T to build the constitutive relationship.

$$G = u - \frac{1}{\rho}\boldsymbol{\sigma} : \boldsymbol{\varepsilon} - sT \tag{1}$$

Here, $\boldsymbol{\varepsilon}$, u and s refer to the total strain tensor, specific internal energy, and specific entropy, respectively. The variable ρ represents the material density and the operator ":" denotes double dot product of tensors. For isotropic polycrystalline SMAs, the Gibbs-free energy is a function of the transformation strain ($\boldsymbol{\varepsilon}^t$), and the martensite volume fraction (ξ) tracks the evolution of the phase transformation. This function can be represented as the sum of the bulk (G_b) and mixing (G_m) energies such that

$$G = G_b + G_m \tag{2}$$

where

$$G_b(\boldsymbol{\sigma}, T, \xi) = -\frac{1}{2\rho}\boldsymbol{\sigma} : \boldsymbol{S} : \boldsymbol{\sigma} - \frac{1}{\rho}\boldsymbol{\sigma} : \boldsymbol{\alpha}(T - T_0) + c\left[(T - T_0) - T\ln\frac{T}{T_0}\right] - s_0 T + u_0 \tag{3}$$

and

$$G_m(\boldsymbol{\sigma}, T, \xi, \boldsymbol{\varepsilon}^t) = -\frac{1}{\rho}\boldsymbol{\sigma} : \boldsymbol{\varepsilon}^t + \frac{1}{\rho}f(\xi) \tag{4}$$

In the above equations, T_0 refers to the reference temperature for thermal expansion, $f(\xi)$ is the transformation hardening function that depicts the transformation energy, and $\boldsymbol{S}$ and $\boldsymbol{\alpha}$ are the compliance tensor and thermal expansion tensor respectively. The transformation strain $\boldsymbol{\varepsilon}^t$ captures

3

the inelastic behavior of the phenomenological model. The compliance tensor $\boldsymbol{S}$ for isotropic SMAs is defined via the compliance modulus S and Poisson's ratio v as follows, where S is the reciprocal of Young's modulus ($S = 1/E$).

$$
\boldsymbol{S} = S : \mathfrak{C}
$$

$$
\mathfrak{C} = \begin{bmatrix}
1 & -v & -v & 0 & 0 & 0 \\
-v & 1 & -v & 0 & 0 & 0 \\
-v & -v & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 2(1+v) & 0 & 0 \\
0 & 0 & 0 & 0 & 2(1+v) & 0 \\
0 & 0 & 0 & 0 & 0 & 2(1+v)
\end{bmatrix}
\tag{5}
$$

The thermal expansion tensor $\boldsymbol{\alpha}$, which leads to purely hydrostatic thermal expansion, is defined by the thermal expansion parameter α in Equation 6. In the equation,

$$
\boldsymbol{\alpha} = \alpha \cdot \mathrm{diag} \begin{pmatrix} 1, & 1, & 1, & 0, & 0, & 0 \end{pmatrix}
\tag{6}
$$

In addition, the symbols c, s_0, and u_0 in Equation 3 represent the effective specific heat, effective specific entropy, and effective specific internal energy, respectively. These parameters are functions of their values in the martensite (M) and austenite (A) phases, and the martensite volume fraction ξ, shown in Equation 7. [1]

$$
\begin{aligned}
S &= S^A + \xi(S^M - S^A) = S^A + \xi \Delta S \\
\alpha &= \alpha^A = \alpha^M \\
c &= c^A = c^M \\
s_0 &= s_0^A + \xi(s_0^M - s_0^A) = s_0^A + \xi \Delta s_0 \\
u_0 &= u_0^A + \xi(u_0^M - u_0^A) = u_0^A + \xi \Delta u_0
\end{aligned}
\tag{7}
$$

The admissible total strain tensor of SMAs $\boldsymbol{\varepsilon}$ should follow the second law of thermodynamics [39].

$$
\frac{1}{\rho}\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - (\dot{u} + \dot{s}T) \geq 0
\tag{8}
$$

where ($\dot{\ }$) denotes differentiation with respect to time.

By substituting Equation 1 and Equation 2 into the second law of thermodynamics via the Coleman Noll procedure [40], the definition of the total strain tensor of SMAs $\boldsymbol{\varepsilon}$ can be decomposed into three parts.

$$
\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^e + \boldsymbol{\varepsilon}^{th} + \boldsymbol{\varepsilon}^t
\tag{9}
$$

where $\boldsymbol{\varepsilon}^e = \boldsymbol{S} : \boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}^{th} = \boldsymbol{\alpha}(T - T_0)$ refer to the elastic strain and pure thermal expansion strain, respectively. Furthermore, the second law, also referred as the Clausius-Planck inequality can be simplified to

$$
\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}}^t - \rho \frac{\partial G}{\partial \xi} \dot{\xi} \geq 0
\tag{10}
$$

---

[1]Note that it is generally assumed that the thermal expansion parameter α and specific heat c of SMAs in the martensite and austenite phases are the same [5], as their relative differences could be smaller than 5%. Further, the pure thermal expansion effect contributes only a small portion of the total strain, compared with the inelastic deformation. However, this work is adaptable to assumptions of different thermal expansion parameters and specific heat in the martensite and austenite phases.

4

To further simplify the inequality, an evolution relation (flow rule) in Equation 11 is proposed to link the transformation strain ε^t and the martensite volume fraction ξ through the transformation tensor $\boldsymbol{\Lambda}$. In this paper, we use the relationship proposed by Boyd and Lagoudas [15].

$$\dot{\varepsilon}^t = \boldsymbol{\Lambda}\dot{\xi}$$

$$\boldsymbol{\Lambda} = \begin{cases} \frac{3}{2}H\frac{\boldsymbol{\sigma}_s}{\sigma_s^{eff}} & \dot{\xi} > 0 \\ H\frac{\boldsymbol{\varepsilon}_{t-r}}{\varepsilon_{t-r}^{eff}} & \dot{\xi} < 0 \end{cases} \tag{11}$$

where H is the maximum transformation strain. For the forward transformation (i.e. $\dot{\xi} > 0$), $\boldsymbol{\sigma}_s$ is the deviatoric stress tensor and σ_s^{eff} is its associated effective (von Mises) stress. In the reverse transformation (i.e. $\dot{\xi} < 0$), $\boldsymbol{\varepsilon}_{t-r}$ is the transformation strain tensor at the reversal point, and ε_{t-r}^{eff} is its associated effective strain. By substituting the flow rule into Equation 10, the Clausius-Planck inequality can be rewritten as

$$\Pi\dot{\xi} \geq 0, \tag{12}$$

where

$$\Pi(\boldsymbol{\sigma}, T, \xi) = \boldsymbol{\sigma} : \Lambda + \frac{1}{2}\boldsymbol{\sigma} : \Delta\boldsymbol{S} : \boldsymbol{\sigma} + \boldsymbol{\sigma} : \Delta\alpha(T - T_0) \\ - \rho\Delta c\left[(T - T_0) - T\ln\frac{T}{T_0}\right] + \rho\Delta s_0 T - \rho\Delta u_0 - \frac{\partial f(\xi)}{\partial\xi} \tag{13}$$

The Clausius-Planck inequality stated in Equation 12 must be satisfied for all admissible thermomechanical loading paths $\dot{\xi}$. These thermomechanical loading paths can be described by a Kuhn-Tucker condition in Equation 14, where $\Phi = |\Pi| - Y$ is a type of yield function and Y is a type of yield strength (transformation threshold) determined by the transformation hardening function [5]. There are two possibilities when $\Phi = 0$. If $\Pi - Y = 0$ then to satisfy Equation 12, $\dot{\xi} > 0$, indicating the forward transformation; otherwise if $-\Pi - Y = 0$ then $\dot{\xi} < 0$, indicating the inverse transformation. For all other cases of thermoelastic behaviors, i.e. for $\Phi < 0$, we have $\dot{\xi} = 0$. [5]

$$\Phi\begin{cases} = \Pi - Y = 0 & \dot{\xi} > 0 \ (\text{A} \to \text{M}) \\ = -\Pi - Y = 0 & \dot{\xi} < 0 \ (\text{M} \to \text{A}) \\ < 0 & \dot{\xi} = 0 \end{cases} \tag{14}$$

In addition, a consistency condition is defined to ensure the stress and temperature remain continuous at the transformation surface throughout the evolution process. [41]

$$\dot{\Phi} = \frac{\partial\Phi}{\partial\boldsymbol{\sigma}} : \dot{\boldsymbol{\sigma}} + \frac{\partial\Phi}{\partial T} : \dot{T} + \frac{\partial\Phi}{\partial\xi} : \dot{\xi} \tag{15}$$

By satisfying the Kuhn-Tucker condition and the consistency condition, the inelastic constitutive relationship can be obtained. In equation 16, $\otimes$ refers to the tensor product operator, $\mathfrak{L}$ and $\boldsymbol{\Theta}$ are the continuum tangent stiffness and tangent moduli. The symbols $\partial_{\boldsymbol{\sigma}}$ and ∂_{ξ} denote partial differentiation with respect to the stress tensor $\boldsymbol{\sigma}$ and the martensite volume fraction ξ, respectively.

$$\mathrm{d}\boldsymbol{\sigma} = \mathfrak{L} : \mathrm{d}\varepsilon + \boldsymbol{\Theta} : \mathrm{d}T$$

$$\mathfrak{L} = \begin{cases} \boldsymbol{S}^{-1} - \frac{\boldsymbol{S}^{-1}:\partial_{\boldsymbol{\sigma}}\Phi \otimes \boldsymbol{S}^{-1}:\partial_{\boldsymbol{\sigma}}\Phi}{\partial_{\boldsymbol{\sigma}}\Phi:\boldsymbol{S}^{-1}:\partial_{\boldsymbol{\sigma}}\Phi - \partial_{\xi}\Phi} & \dot{\xi} > 0 \\ \boldsymbol{S}^{-1} - \frac{\boldsymbol{S}^{-1}:\partial_{\boldsymbol{\sigma}}\Phi \otimes \boldsymbol{S}^{-1}:\partial_{\boldsymbol{\sigma}}\Phi}{\partial_{\boldsymbol{\sigma}}\Phi:\boldsymbol{S}^{-1}:\partial_{\boldsymbol{\sigma}}\Phi + \partial_{\xi}\Phi} & \dot{\xi} < 0 \end{cases} \tag{16}$$

$$\boldsymbol{\Theta} = \begin{cases} -\mathfrak{L} : \boldsymbol{\alpha} - \partial_T\Phi\frac{\boldsymbol{S}^{-1}:\partial_{\boldsymbol{\sigma}}\Phi}{\partial_{\boldsymbol{\sigma}}\Phi:\boldsymbol{S}^{-1}:\partial_{\boldsymbol{\sigma}}\Phi - \partial_{\xi}\Phi} & \dot{\xi} > 0 \\ -\mathfrak{L} : \boldsymbol{\alpha} - \partial_T\Phi\frac{\boldsymbol{S}^{-1}:\partial_{\boldsymbol{\sigma}}\Phi}{\partial_{\boldsymbol{\sigma}}\Phi:\boldsymbol{S}^{-1}:\partial_{\boldsymbol{\sigma}}\Phi + \partial_{\xi}\Phi} & \dot{\xi} < 0 \end{cases}$$

5

2.2. Thermomechanical Coupled Model of SMAs

For the constitutive model, we consider the coupling effect between inelasticity and transient thermal conduction of SMAs. The necessity of including transient thermal conduction in the framework is related to the accuracy of the finite element model used in the optimization process. First, the phenomenological model (cf. Equation 14) is a transient, path-dependent functional of the derivatives of temperature with respect to time. As a result, a steady-heat model, which assumes the derivative of temperature is uniformly zero, is not suitable for our framework. Moreover, in the context of multimaterial design for nonlinear thermal actuation, the cooling and heating rate of the different materials may vary significantly for each material candidate due to their different thermal conductivity. Such differences will influence the transient thermal behavior of the materials.

The governing equations of the model are stated in Equation 17, with the natural boundary conditions specified in Figure 1.

$$
\begin{aligned}
\text{Global level}: \quad &\text{Force equilibrium} && \int_\Omega \varepsilon(\delta\boldsymbol{d})\boldsymbol{\sigma}(\boldsymbol{d},\xi)\mathrm{d}\Omega - \int_{\Gamma_{\boldsymbol{p}}} \delta\boldsymbol{d}\boldsymbol{p}\mathrm{d}\Gamma_{\boldsymbol{p}} = 0 \quad \forall\, \delta\boldsymbol{d} \\
&\text{Thermal conduction} && \int_\Omega [\tfrac{1}{2}\bigtriangledown\boldsymbol{T}:\mathbb{K}:\bigtriangledown\boldsymbol{T} - (\boldsymbol{Q}-\rho c\tfrac{\partial\boldsymbol{T}}{\partial t})\boldsymbol{T}]\mathrm{d}\Omega \\
& && \quad - \int_{\Gamma_{\boldsymbol{q}}}\boldsymbol{q}\boldsymbol{T}\mathrm{d}\Gamma_{\boldsymbol{q}} = 0 \\
\text{Local level}: \quad &\text{KKT}-\text{condition} && \Phi \le 0 \text{ in } \Omega \\
& && \dot{\xi}\Phi = 0 \text{ in } \Omega \\
&\text{Consistency} && \dot{\Phi} = \tfrac{\partial\Phi}{\partial\boldsymbol{\sigma}}:\dot{\boldsymbol{\sigma}} + \tfrac{\partial\Phi}{\partial\boldsymbol{T}}:\dot{\boldsymbol{T}} + \tfrac{\partial\Phi}{\partial\xi}:\dot{\xi} \text{ in } \Omega \\
&\text{Flow rule} && \dot{\boldsymbol{\varepsilon}}^t = \boldsymbol{\Lambda}\dot{\xi}
\end{aligned}
\tag{17}
$$

where $\boldsymbol{d}$ and $\boldsymbol{T}$ refer to the displacement and temperature fields, respectively. $\mathbb{K} = \mathrm{diag}(\kappa_x, \kappa_y, \kappa_z)$ is the conduction matrix and κ refers to the thermal conductivity. Q is the heat source/sink. The parameter t refers to the time step and the symbol $\bigtriangledown$ represents the gradient operator.
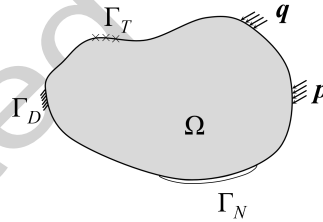
Figure 1: Example representation of the boundary conditions of the coupled thermomechanical problem

On the boundary of the design domain Ω, Γ_D and Γ_T are Dirichlet boundary conditions referring to region with fixed displacement vector and temperature fields, respectively. Γ_N is the Neumann boundary of isolated thermal conduction. The symbols $\boldsymbol{p}$ and $\boldsymbol{q}$ represent Neumann boundary conditions of prescribed force and heat flux, respectively. Note that the thermomechanical problem contains one-directional coupling. The temperature field $\boldsymbol{T}$ is solely updated via solving the transient thermal conduction problem, which does not depend on the solution of the nonlinear mechanical problem used to update the displacement field $\boldsymbol{d}$.

2.3. Finite Element Analysis

Discretizing the system via a Backward-Euler scheme with $\frac{\partial\boldsymbol{T}_n}{\partial t} = \frac{\boldsymbol{T}_n - \boldsymbol{T}_{n-1}}{\Delta t}$, Equation 17 is further expressed through the finite element method as three sets of residuals $\boldsymbol{R}$, $\boldsymbol{G}$ and $\boldsymbol{H}$.

$$
\text{Global level}: \boldsymbol{R}_{n+1} = \bigwedge_{\text{el}} [\sum_{\mathfrak{G}^\Omega}(w\boldsymbol{B}_{\mathfrak{G}^\Omega}^{\mathrm{T}}\boldsymbol{\sigma}_{\mathfrak{G}^\Omega,n+1}) - \sum_{\mathfrak{G}^\Gamma}(\varpi\boldsymbol{p}_{\mathfrak{G}^\Gamma,n+1}\boldsymbol{N}_{\mathfrak{G}^\Gamma})]\det\boldsymbol{J} = \boldsymbol{0}
$$

6

$$G_{n+1} = (K_c + \frac{C}{\Delta t})T_{n+1} - \frac{C}{\Delta t}T_n - q_{n+1}^c = 0 \tag{18}$$

$$\text{Local level}: \begin{cases} H_{n+1} = \begin{cases} H_\xi = \Phi(\sigma_{n+1}, T_{n+1}, \xi_{n+1}) \\ H_{\varepsilon^t} = \varepsilon_n^t + \Lambda(\xi_{n+1} - \xi_n) - \varepsilon_{n+1}^t \\ H_S = S_n + \Delta S(\xi_{n+1} - \xi_n) - S_{n+1} \\ H_\sigma = S_{n+1}^{-1} : [\varepsilon_{n+1} - \alpha(T_{n+1} - T_0) - \varepsilon_{n+1}^t] - \sigma_{n+1} \end{cases} = 0 \quad \dot{\xi} \neq 0 \\ H_{n+1} = S^{-1} : [(\varepsilon_{n+1} - \alpha(T_{n+1} - T_0) - \varepsilon^t) - \sigma_{n+1} = 0 \qquad\qquad \dot{\xi} = 0 \end{cases}$$

where the global force residual R and thermal loading residual G contain the nodal information of displacement and temperature fields at global degrees of freedom, respectively. The symbol $\bigwedge$ denotes the assembly operator for assembling global matrices, and its corresponding subscript $_{\text{el}}$, refers to one discrete element. The superscript $^{\text{T}}$ refers to the transpose operator. The local residual H defines the constitutive relationship of SMAs at the local level, i.e. Gauss points. Note that the total strain ε_{n+1} can be expressed as Bd_{n+1}. In addition, the thermal coefficient matrices and loading vector in Equation 18 are defined as

$$K_c = \bigwedge_{\text{el}} \sum_{\mathfrak{G}^\Omega} w B_{\mathfrak{G}^\Omega}^{\text{T}} \mathbb{K} B_{\mathfrak{G}^\Omega} \det J$$

$$C = \bigwedge_{\text{el}} \sum_{\mathfrak{G}^\Omega} w N_{\mathfrak{G}^\Omega}^{\text{T}} \rho c N_{\mathfrak{G}^\Omega} \det J \tag{19}$$

$$q_{n+1}^c = \bigwedge_{\text{el}} (\sum_{\mathfrak{G}^\Omega} w Q_{\mathfrak{G}^\Omega, n+1} N_{\mathfrak{G}^\Omega} + \sum_{\mathfrak{G}^\Gamma} \varpi q_{\mathfrak{G}^\Gamma, n+1} N_{\mathfrak{G}^\Gamma}) \det J$$

In Equations 18 and 19, $\mathfrak{G}^\Omega$, $\mathfrak{G}^\Gamma$ are Gauss points for integration on the design domain Ω and the boundary Γ, respectively, with w and ϖ serving as the corresponding Gaussian weights. $\det J$ is the determinant of the Jacobian matrix. The matrix B contains the spatial derivatives of the shape functions, N. The subscripts of B and N indicate evaluation at the corresponding Gauss points. For the purpose of sensitivity analysis, we further define $\widetilde{K}_c = K_c + \frac{C}{\Delta t}$ and $\widetilde{C} = \frac{C}{\Delta t}$.

Then, we uncouple the nonlinear system shown in Equation 18 through a global-local approach,

$$\begin{aligned} R(u, \nu(u, \omega), \omega) &= 0 \\ H(u, \nu(u, \omega), \omega) &= 0 \\ G(\omega) &= 0 \end{aligned} \tag{20}$$

This system can be solved numerically via a nested Newton-Raphson iteration (return mapping algorithm [42] or a cross-iterative Newton-Raphson procedure [43]. The tensors u, ν and ω are thermomechanical state variables. Specifically, the local mechanical state variables $\nu = [\xi, \varepsilon^t, S, \sigma]$, is a function of the global mechanical state variable $u = [F^c, d^f]$, and the global thermal state variable $\omega = [T^d, q^p]$. Furthermore, each state variable is partitioned based upon which degrees of freedom are free and fixed. The symbol d^f is the displacement vector at free degrees of mechanical freedom and F^c is the unknown applied force at constrained degrees of mechanical freedom. Also, T^d refers to the temperature vector at free degrees of thermal freedom, and q^p stands for the unknown heat flux at prescribed degrees of thermal freedom. Figure 2 contains flowcharts illustrating the flow of data during the finite element analysis procedure at time step $n+1$ based on the return mapping (left) and parallel projection (right) algorithms.

In Figure 2, k and l are iteration counters of the Newton-Raphson procedures. Note that the transient thermal conduction problem at each time step is directly solved without an additional

7

(a) Return Mapping   (b) Parallel Projection
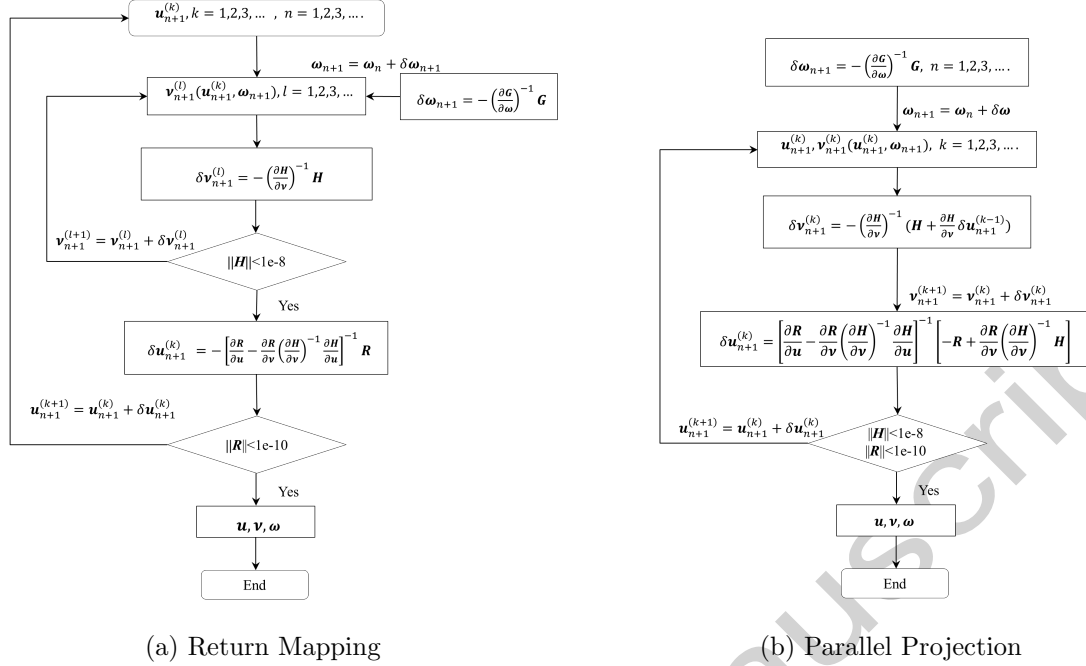
Figure 2: Flow charts illustrating the return mapping and parallel projection algorithms for solving inelasticity analysis coupled with transient thermal conduction

iteration procedure. Note that the time step Δt is implicitly related to the temperature increment from T_n to T_{n+1}. Note that due to that the constitutive law of SMAs is highly nonlinear, the unconditional stability of the backward Euler method is not always preserved. The choice of Δt should lead to a temperature increment with an absolute value generally no larger than 0.6K, which is around the maximum allowable temperature step for both the return mapping and parallel projection algorithms [43]. For numerical details of the aforementioned algorithms, please refer to the authors' previous research [43].

Figure 3 shows samples of phase diagrams of SMAs, which depict the variation of the transformation temperature and transformation stress of SMAs based on different thermal and mechanical conditions. With the proposed thermomechanical coupled model, the complex physical interaction between the transformation stress and transformation temperature is intrinsically embedded into the finite element analysis.

2.4. General Topology Optimization Problem

In this paper, we seek an optimization framework that enables us to tailor and maximally leverage the unique behaviors of SMAs in a multi-material design. To begin, we give the definition of a general topology optimization problem. For a design domain Ξ subject to Dirichlet Γ_{Drt} and Neumann boundary Γ_{Neu} conditions (Figure 4), topology optimization provides an optimal distribution of materials Ω via interpreting the multi-physics performance of the structure. Such a problem is solved by minimizing an objective functional f_{obj}, while satisfying a set of equality and inequality constraints c_i and c_j (Equation 21). Here, $\boldsymbol{\eta}$ is the design variable used to represent the presence of material at a given point within the design domain.

8

(a) Transformation temperature under TWSME

(b) Transformation stress under superelasticity
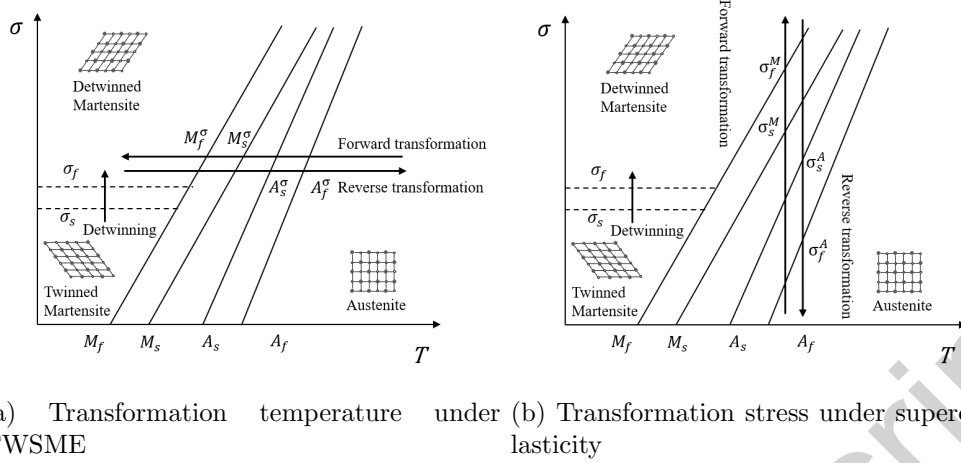
Figure 3: Example phase diagrams
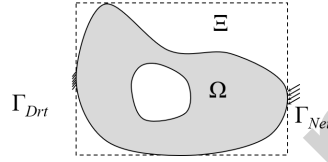
Figure 4: Illustration of topology optimization

$$
\begin{aligned}
\min_{\partial\Omega} \quad & f_{obj}(\boldsymbol{\eta}, \Omega) \\
s.t \quad & c_i(\boldsymbol{\eta}) = 0, \ i = 1, 2, ..., n_i \\
& c_j(\boldsymbol{\eta}) \geq 0, \ j = 1, 2, ..., n_j
\end{aligned}
\tag{21}
$$

2.5. Multi-material Design Parameterization

Herein, we select the multi-material structure to be composed by SMAs for actuation, and inactive materials for programming the direction of actuation. Note that the inactive materials do not have transformation behaviors but only behave thermoelastically. To optimize the transient thermomechanial behavior of the structure, we characterize the design materials by their Young's modulus E, thermal conductivity κ and mass density ρ. For example, the effective Young's modulus of a given element, i, is evaluated as a weighted sum of the actual Young's modulus, $E^{(j)}$, of the candidate materials. [2]

$$
\begin{aligned}
E_i &= \sum_{j=1}^{m} \mu_i^{(j)} E^{(j)} \\
\kappa_i &= \sum_{j=1}^{m} \mu_i^{(j)} \kappa^{(j)} \\
\rho_i &= \sum_{j=1}^{m} \mu_i^{(j)} \rho^{(j)}
\end{aligned}
\tag{22}
$$

---

[2]For simplicity, we assume that all the material candidates have the same thermal expansion coefficient(α), specific heat (c) and Poisson ratio v.

9

where j refers to the numbering of the material candidates.

For a topology optimization problem considering two materials candidates (solid plus void), the *Solid Isotropic Material with Penalisation* (SIMP) method is adopted [44]]. As shown in Equation 23, the SIMP method has only one design parameter $r \in [0, 1]$ to represent the relative volume fraction μ of each material candidate within a given element.

$$\mu^{(1)} = r^P$$
$$\mu^{(2)} = (1 - r)^P$$
(23)

where P is the penalization constant used to avoid intermediate material states.

In terms of the multi-material case, the material distribution is represented parametrically using the *Shape Functions with Penalization* (SFP) method [45, 46]. In the case of four material candidates for example, the relative volume fraction of a given candidate material, $\mu^{(j)}$, $j = 1, 2, 3, 4$, can be expressed as 24.

$$\mu^{(1)} = r_1^P r_2^P$$
$$\mu^{(2)} = r_1^P (1 - r_2)^P$$
$$\mu^{(3)} = (1 - r_1)^P r_2^P$$
$$\mu^{(4)} = (1 - r_1)^P (1 - r_2)^P$$
(24)

where $r_1, r_2 \in [0, 1]$ are independent design parameters, and P is the penalization factor. Figure 5 presents the activation functions of the SFP method in the $r_1 - r_2$ plane, in the case of $P = 1$. When $P > 1$, the activation functions will become convex and disfavor intermediate values. It can be observed that the SFP method is similar to the SIMP method in that a higher penalty value will encourage the removal of intermediate density (i.e. hybrid) materials within an element. Note also that for the four-material case, the SFP method saves computation cost by requiring only two design parameters per element, while the SIMP method requires at least three design parameters [45].
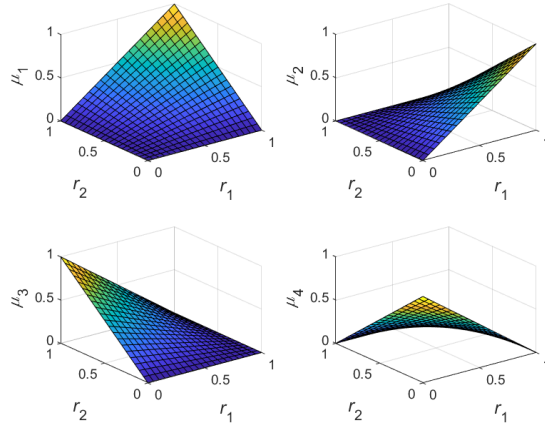
Figure 5: Activation functions of the SFP method for a four-material problem, $P = 1$.

Now, we discuss the evaluation of transformation status of the SMA-based multi-material structure at the elemental level. We compute the global internal force using the following relationship

$$f_{\text{int}} = \bigwedge_{\text{el}} \sum_{\mathfrak{G}} \boldsymbol{\sigma}_{\mathfrak{G}} = \bigwedge_{\text{el}} \sum_{\mathfrak{G}} (\boldsymbol{\sigma}_{\mathfrak{G}}^{\text{elas}} + \boldsymbol{\sigma}_{\mathfrak{G}}^{\text{inelas}})$$
(25)

10

$$= \bigwedge_{\text{el}} \sum_{\mathfrak{G}} \left\{ \sum_{j=1, j\neq k}^{m} \mu_{\text{el}}^{(j)} E^{(j)} : \mathfrak{C} : [\boldsymbol{\varepsilon}_{\mathfrak{G}} - \boldsymbol{\alpha}(T - T_0)] + \mu_{\text{el}}^{(k)} E^{*(k)}_{\mathfrak{G}} : \mathfrak{C} : [\boldsymbol{\varepsilon}_{\mathfrak{G}} - \boldsymbol{\alpha}(T - T_0) - \varepsilon^t] \right\}$$

where k is the index of the SMA material, and $E^* = 1/S$ is the evolving Young's modulus of the SMA at each Gauss point $\mathfrak{G}$ due to transformation. Noticeably, we consider that the stress tensor of a multimaterial SMA-based structure is composed of the elastic stress $\boldsymbol{\sigma}^{\text{elas}}$ caused by the inactive materials and the inelastic stress $\boldsymbol{\sigma}^{\text{inelas}}$ caused by SMAs. Consequently, we only evaluate the transformation effects of the SMA material.[3] Note that the constitutive law of the inactive materials does not require to be calculated in a local residual form separately, i.e. $\boldsymbol{H}_n = E : \mathfrak{C} : (\boldsymbol{\varepsilon}_n - \boldsymbol{\alpha}(T_n - T_0)) - \boldsymbol{\sigma}_n = 0$. Instead, the constitutive law can be directly incorporated into the global residual, such that

$$\begin{aligned}
\boldsymbol{R}_n &= \bigwedge_{\text{el}} \left[\sum_{\mathfrak{G}^\Omega} (w \boldsymbol{B}_{\mathfrak{G}^\Omega}^{\text{T}} \boldsymbol{\sigma}_{\mathfrak{G}^\Omega, n}) - \sum_{\mathfrak{G}^\Gamma} (\varpi \boldsymbol{p}_{\mathfrak{G}^\Gamma, n} \boldsymbol{N}_{\mathfrak{G}^\Gamma}) \right] \det \boldsymbol{J} \\
&= \bigwedge_{\text{el}} \left\{ \sum_{\mathfrak{G}^\Omega} [w \boldsymbol{B}_{\mathfrak{G}^\Omega}^{\text{T}} (\boldsymbol{\sigma}_{\mathfrak{G}^\Omega, n}^{\text{elas}} + \boldsymbol{\sigma}_{\mathfrak{G}^\Omega, n}^{\text{inelas}})] - \sum_{\mathfrak{G}^\Gamma} (\varpi \boldsymbol{p}_{\mathfrak{G}^\Gamma, n} \boldsymbol{N}_{\mathfrak{G}^\Gamma}) \right\} \det \boldsymbol{J} \\
&= \bigwedge_{\text{el}} \left\{ \sum_{\mathfrak{G}^\Omega} [w \boldsymbol{B}_{\mathfrak{G}^\Omega}^{\text{T}} (\sum_{j=1, j\neq k}^{m} \mu_{\text{el}}^{(j)} E^{(j)} : \mathfrak{C} : [\boldsymbol{\varepsilon}_{\mathfrak{G}} - \boldsymbol{\alpha}(T - T_0)] + \boldsymbol{\sigma}_{\mathfrak{G}^\Omega, n}^{\text{inelas}})] - \sum_{\mathfrak{G}^\Gamma} (\varpi \boldsymbol{p}_{\mathfrak{G}^\Gamma, n} \boldsymbol{N}_{\mathfrak{G}^\Gamma}) \right\} \det \boldsymbol{J}
\end{aligned} \tag{26}$$

2.6. Sensitivity Analysis

2.6.1. The Transient Adjoint Sensitivity Formulation

Building upon on the authors' previous work on sensitivity analysis of SMA-based structures with a uniform temperature distribution [38], we again use a transient, path-dependent adjoint method. Considering the coupled transient thermal conduction problem, the global thermal residual G must be included into Lagrangian functional in addition to the mechanical residuals $\boldsymbol{R}$ and $\boldsymbol{H}$,

$$\begin{aligned}
\Pi =\, & f_{int}(\boldsymbol{u}_{N_t}(\boldsymbol{r}), \boldsymbol{\nu}_{N_t}(\boldsymbol{r}), \boldsymbol{r}) + \sum_{n=1}^{N_t} \boldsymbol{\lambda}_n^{\text{T}} \boldsymbol{R}_n(\boldsymbol{u}_n(\boldsymbol{r}), \boldsymbol{u}_{n-1}(\boldsymbol{r}), \boldsymbol{\nu}_n(\boldsymbol{r}), \boldsymbol{\nu}_{n-1}(\boldsymbol{r}), \boldsymbol{\omega}_n(\boldsymbol{r}), \boldsymbol{\omega}_{n-1}(\boldsymbol{r}), \boldsymbol{r}) \\
& + \sum_{n=1}^{N_t} \left(\bigwedge_{\text{el}} \sum_{\mathfrak{G}} \boldsymbol{\gamma}_{\mathfrak{G}, n}^{\text{T}} \boldsymbol{H}_{\mathfrak{G}, n}(\boldsymbol{u}_n(\boldsymbol{r}), \boldsymbol{u}_{n-1}(\boldsymbol{r}), \boldsymbol{\nu}_{\mathfrak{G}, n}(\boldsymbol{r}), \boldsymbol{\nu}_{\mathfrak{G}, n-1}(\boldsymbol{r}), \boldsymbol{\omega}_n(\boldsymbol{r}), \boldsymbol{\omega}_{n-1}(\boldsymbol{r}), \boldsymbol{r}) \right) \\
& + \sum_{n=1}^{N_t} \boldsymbol{\psi}_n^{\text{T}} \boldsymbol{G}_n(\boldsymbol{u}_n(\boldsymbol{r}), \boldsymbol{u}_{n-1}(\boldsymbol{r}), \boldsymbol{\nu}_n(\boldsymbol{r}), \boldsymbol{\nu}_{n-1}(\boldsymbol{r}), \boldsymbol{\omega}_n(\boldsymbol{r}), \boldsymbol{\omega}_{n-1}(\boldsymbol{r}), \boldsymbol{r})
\end{aligned} \tag{27}$$

where f_{int} represents an arbitrary function of interest (either an objective or constraint functions).

In Equation 27, the coefficient vectors $\boldsymbol{\lambda}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\psi}$ are free parameters whose values can be chosen to maximize computational expediency. Note that all of the local adjoint vectors $\boldsymbol{\gamma}$ are evaluated at the Gauss points, since the local constitutive law $\boldsymbol{H}$ is evaluated at Gauss points.

---

[3]Note that one may also treat the multi-material composite within each element as a whole unit having transformation behavior, i.e. $f_{\text{int}} = \bigwedge_{\text{el}} \sum_{\mathfrak{G}} \sum_{j=1}^{m} \mu_{\text{el}}^{(j)} E^{(j)} [\boldsymbol{\varepsilon}^{\mathfrak{G}} - \boldsymbol{\alpha}(T - T_0) - \varepsilon^t]$. However, extremely strong assumptions on physical properties of the composite are needed so that the transformation temperature of the composite unit is identical to that of the SMA, or the transformation temperature of the composite follows a proposed interpolation rule. One also needs to make the assumption that the transformation behaviors of the physical properties of the composites still follow Equation 7. To avoid such strong assumptions, we suggest decomposing the stress tensor and evaluating the transformation behavior only for the portion of each element that contains SMAs.

11

Since all residuals $\boldsymbol{R}$, $\boldsymbol{H}$ and $\boldsymbol{G}$ are zero vectors, the equality $\frac{\mathrm{d}f_{int}}{\mathrm{d}\boldsymbol{r}} = \frac{\mathrm{d}\Pi}{\mathrm{d}\boldsymbol{r}}$ holds for all values of the vectors $\boldsymbol{\lambda}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\psi}$. Taking the first derivative of Π with respect to the design variable vector $\boldsymbol{r}$ and applying the chain rule, we obtain

$$
\begin{aligned}
\frac{\mathrm{d}\Pi}{\mathrm{d}\boldsymbol{r}} = \frac{\partial f_{int}}{\partial \boldsymbol{r}} &+ \sum_{n+1}^{N_t}\left(\frac{\partial f_{int}}{\partial \boldsymbol{u}_n}\frac{\mathrm{d}\boldsymbol{u}_n}{\mathrm{d}\boldsymbol{r}} + \sum_{\mathfrak{G}}\frac{\partial f_{int}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}}\frac{\mathrm{d}\boldsymbol{\nu}_{\mathfrak{G},n}}{\mathrm{d}\boldsymbol{r}} + \frac{\partial f_{int}}{\partial \boldsymbol{\omega}_n}\frac{\mathrm{d}\boldsymbol{\omega}_n}{\mathrm{d}\boldsymbol{r}}\right) + \sum_{n=1}^{N_t}\boldsymbol{\lambda}_n^{\mathrm{T}}\left[\frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{r}} + \frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{u}_n}\frac{\mathrm{d}\boldsymbol{u}_n}{\mathrm{d}\boldsymbol{r}} + \right. \\
&\frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{u}_{n-1}}\frac{\mathrm{d}\boldsymbol{u}_{n-1}}{\mathrm{d}\boldsymbol{r}} + \bigwedge_{\mathrm{el}}\sum_{\mathfrak{G}}\left(\frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}}\frac{\mathrm{d}\boldsymbol{\nu}_{\mathfrak{G},n}}{\mathrm{d}\boldsymbol{r}} + \frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\nu}_{\mathfrak{G},n-1}}\frac{\mathrm{d}\boldsymbol{\nu}_{\mathfrak{G},n-1}}{\mathrm{d}\boldsymbol{r}}\right) + \left.\frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\omega}_n}\frac{\mathrm{d}\boldsymbol{\omega}_n}{\mathrm{d}\boldsymbol{r}} + \frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\omega}_{n-1}}\frac{\mathrm{d}\boldsymbol{\omega}_{n-1}}{\mathrm{d}\boldsymbol{r}}\right] \\
&+ \sum_{n=1}^{N_t}\left[\bigwedge_{\mathrm{el}}\sum_{\mathfrak{G}}\boldsymbol{\gamma}_{\mathfrak{G},n}^{\mathrm{T}}\left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{r}} + \frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{u}_n}\frac{\mathrm{d}\boldsymbol{u}_n}{\mathrm{d}\boldsymbol{r}} + \frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{u}_{n-1}}\frac{\mathrm{d}\boldsymbol{u}_{n-1}}{\mathrm{d}\boldsymbol{r}} + \frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}}\frac{\mathrm{d}\boldsymbol{\nu}_{\mathfrak{G},n}}{\mathrm{d}\boldsymbol{r}}\right.\right. \qquad (28) \\
&\left.\left.+\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n-1}}\frac{\mathrm{d}\boldsymbol{\nu}_{\mathfrak{G},n-1}}{\mathrm{d}\boldsymbol{r}} + \frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\omega}_n}\frac{\mathrm{d}\boldsymbol{\omega}_n}{\mathrm{d}\boldsymbol{r}} + \frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\omega}_{n-1}}\frac{\mathrm{d}\boldsymbol{\omega}_{n-1}}{\mathrm{d}\boldsymbol{r}}\right)\right] + \sum_{n=1}^{N_t}\boldsymbol{\psi}_n^{\mathrm{T}}\left[\frac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{r}} + \frac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{u}_n}\frac{\mathrm{d}\boldsymbol{u}_n}{\mathrm{d}\boldsymbol{r}} + \right. \\
&\frac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{u}_{n-1}}\frac{\mathrm{d}\boldsymbol{u}_{n-1}}{\mathrm{d}\boldsymbol{r}} + \bigwedge_{\mathrm{el}}\sum_{\mathfrak{G}}\left(\frac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}}\frac{\mathrm{d}\boldsymbol{\nu}_{\mathfrak{G},n}}{\mathrm{d}\boldsymbol{r}} + \frac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{\nu}_{\mathrm{G},n-1}}\frac{\mathrm{d}\boldsymbol{\nu}_{\mathfrak{G},n-1}}{\mathrm{d}\boldsymbol{r}}\right) + \left.\frac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{\omega}_n}\frac{\mathrm{d}\boldsymbol{\omega}_n}{\mathrm{d}\boldsymbol{r}} + \frac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{\omega}_{n-1}}\frac{\mathrm{d}\boldsymbol{\omega}_{n-1}}{\mathrm{d}\boldsymbol{r}}\right]
\end{aligned}
$$

where $\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{r}}$ corresponds to implicit derivatives and $\frac{\partial}{\partial \boldsymbol{r}}$ are explicit derivatives. Note that the implicit derivatives ($\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{r}}$), capture implicit dependence of the state variables $\boldsymbol{u}$, $\boldsymbol{\nu}$ and $\boldsymbol{\omega}$ with respect to the design variables due to the solution of the residual equations. For this reason, the implicit derivatives are expensive to compute. Hence, the adjoint formulation is used to find a solution for the free parameters, $\boldsymbol{\lambda}_n$, $\boldsymbol{\gamma}_{\mathfrak{G},n}$ and $\boldsymbol{\psi}_n$, that causes all implicit terms at every transient time step n to vanish. Hence, the sum of all terms containing implicit derivatives are set to zero in the following system of equations.

$$
n = N_t:\begin{cases}
(\frac{\partial f_{int}}{\partial \boldsymbol{u}_{N_t}} + \boldsymbol{\lambda}_{N_t}^{\mathrm{T}}\frac{\partial \boldsymbol{R}_{N_t}}{\partial \boldsymbol{u}_{N_t}} + \bigwedge_{\mathrm{el}}\sum_{\mathfrak{G}}\boldsymbol{\gamma}_{\mathfrak{G},N_t}^{\mathrm{T}}\frac{\partial \boldsymbol{H}_{\mathfrak{G},N_t}}{\partial \boldsymbol{u}_{N_t}} + \boldsymbol{\psi}_{N_t}^{\mathrm{T}}\frac{\partial \boldsymbol{G}_{N_t}}{\partial \boldsymbol{u}_{N_t}})\frac{\mathrm{d}\boldsymbol{u}_{N_t}}{\mathrm{d}\boldsymbol{r}} = 0 \\
(\frac{\partial f_{int}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}} + \boldsymbol{\lambda}_{N_t}^{\mathrm{T}}\frac{\partial \boldsymbol{R}_{N_t}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}} + \boldsymbol{\gamma}_{\mathfrak{G},N_t}^{\mathrm{T}}\frac{\partial \boldsymbol{H}_{\mathfrak{G},N_t}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}} + \boldsymbol{\psi}_{N_t}^{\mathrm{T}}\frac{\partial \boldsymbol{G}_{N_t}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}})\frac{\mathrm{d}\boldsymbol{\nu}_{\mathfrak{G},N_t}}{\mathrm{d}\boldsymbol{r}} = 0,\ \forall\ \mathfrak{G} \\
(\frac{\partial f_{int}}{\partial \boldsymbol{\omega}_{N_t}} + \boldsymbol{\lambda}_{N_t}^{\mathrm{T}}\frac{\partial \boldsymbol{R}_{N_t}}{\partial \boldsymbol{\omega}_{N_t}} + \bigwedge_{\mathrm{el}}\sum_{\mathfrak{G}}\boldsymbol{\gamma}_{\mathfrak{G},N_t}^{\mathrm{T}}\frac{\partial \boldsymbol{H}_{\mathfrak{G},N_t}}{\partial \boldsymbol{\omega}_{N_t}} + \boldsymbol{\psi}_{N_t}^{\mathrm{T}}\frac{\partial \boldsymbol{G}_{N_t}}{\partial \boldsymbol{\omega}_{N_t}})\frac{\mathrm{d}\boldsymbol{\omega}_{N_t}}{\mathrm{d}\boldsymbol{r}} = 0
\end{cases} \qquad (29)
$$

$$
n < N_t:\begin{cases}
\left[\boldsymbol{\lambda}_n^{\mathrm{T}}\frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{u}_n} + \boldsymbol{\lambda}_{n+1}^{\mathrm{T}}\frac{\partial \boldsymbol{R}_{n+1}}{\partial \boldsymbol{u}_n} + \bigwedge_{\mathrm{el}}\sum_{\mathfrak{G}}(\boldsymbol{\gamma}_{\mathfrak{G},n}^{\mathrm{T}}\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{u}_n} + \boldsymbol{\gamma}_{\mathfrak{G},n+1}^{\mathrm{T}}\frac{\partial \boldsymbol{H}_{\mathfrak{G},n+1}}{\partial \boldsymbol{u}_n}) + \boldsymbol{\psi}_n^{\mathrm{T}}\frac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{u}_n} + \boldsymbol{\psi}_{n+1}^{\mathrm{T}}\frac{\partial \boldsymbol{G}_{n+1}}{\partial \boldsymbol{u}_n}\right]\frac{\mathrm{d}\boldsymbol{u}_n}{\mathrm{d}\boldsymbol{r}} = 0 \\
(\boldsymbol{\lambda}_n^{\mathrm{T}}\frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\nu}_{\mathrm{G},n}} + \boldsymbol{\lambda}_{n+1}^{\mathrm{T}}\frac{\partial \boldsymbol{R}_{n+1}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} + \boldsymbol{\gamma}_{\mathfrak{G},n}^{\mathrm{T}}\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} + \boldsymbol{\gamma}_{\mathfrak{G},n+1}^{\mathrm{T}}\frac{\partial \boldsymbol{H}_{\mathfrak{G},n+1}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} + \boldsymbol{\psi}_n^{\mathrm{T}}\frac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} + \boldsymbol{\psi}_{n+1}^{\mathrm{T}}\frac{\partial \boldsymbol{G}_{n+1}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}})\frac{\mathrm{d}\boldsymbol{\nu}_{\mathfrak{G},n}}{\mathrm{d}\boldsymbol{r}} = 0,\ \forall\ \mathfrak{G} \\
\left[\boldsymbol{\lambda}_n^{\mathrm{T}}\frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\omega}_n} + \boldsymbol{\lambda}_{n+1}^{\mathrm{T}}\frac{\partial \boldsymbol{R}_{n+1}}{\partial \boldsymbol{\omega}_n} + \bigwedge_{\mathrm{el}}\sum_{\mathfrak{G}}(\boldsymbol{\gamma}_{\mathfrak{G},n}^{\mathrm{T}}\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\omega}_n} + \boldsymbol{\gamma}_{\mathfrak{G},n+1}^{\mathrm{T}}\frac{\partial \boldsymbol{H}_{\mathfrak{G},n+1}}{\partial \boldsymbol{\omega}_n}) + \boldsymbol{\psi}_n^{\mathrm{T}}\frac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{\omega}_n} + \boldsymbol{\psi}_{n+1}^{\mathrm{T}}\frac{\partial \boldsymbol{G}_{n+1}}{\partial \boldsymbol{\omega}_n}\right]\frac{\mathrm{d}\boldsymbol{\omega}_n}{\mathrm{d}\boldsymbol{r}} = 0
\end{cases}
$$

The solution of the above adjoint equations is given in Equation 30, yielding the adjoint vectors. Note that the adjoint vectors of the global thermal residual $\boldsymbol{G}$ are independent of the mechanical state variables $\boldsymbol{u}$ and $\boldsymbol{\nu}$, due to the one-directional coupling within governing equations, i.e. Equation 18. Hence the terms $\frac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{u}_n}$, $\frac{\partial \boldsymbol{G}_{n+1}}{\partial \boldsymbol{u}_n}$, $\frac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{\nu}_n}$ and $\frac{\partial \boldsymbol{G}_{n+1}}{\partial \boldsymbol{\nu}_n}$ are intrinsically zero. Consequently, the adjoint vectors $\boldsymbol{\lambda}_n$ and $\boldsymbol{\gamma}_{\mathrm{G},n}$ only rely on the mechanical residuals $\boldsymbol{R}$ and $\boldsymbol{H}$. Hence, the two adjoint vectors have the same formulas as those obtained for sensitivity analysis of SMAs with uniform temperature distribution [38]. Meanwhile, the adjoint vector $\boldsymbol{\psi}_n$, which is related to transient thermal conduction, is influenced by all of the residuals $\boldsymbol{R}$, $\boldsymbol{H}$ and $\boldsymbol{G}$. Note also that $\frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\omega}_n}$ and $\frac{\partial \boldsymbol{R}_{n+1}}{\partial \boldsymbol{\omega}_n}$ are zero for SMAs, since the coupling between the mechanical and thermal behaviors of SMAs is determined solely via the local residual $\boldsymbol{H}$, i.e. $\boldsymbol{H}_n = \boldsymbol{H}_n(\boldsymbol{\omega}_n, \boldsymbol{\omega}_{n-1})$. However, $\frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\omega}_n}$ remains nonzero

12

for inactive materials (see Equation 26), due to the presence of the thermal expansion term in the global residual $\boldsymbol{R}$.

$$
\begin{aligned}
n = N_t : \begin{cases}
\boldsymbol{\lambda}_{N_t} = -\left[\dfrac{\partial \boldsymbol{R}_{N_t}}{\partial \boldsymbol{u}_{N_t}} - \bigwedge_{\text{el}} \sum_{\mathfrak{G}} \dfrac{\partial \boldsymbol{R}_{N_t}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}} \left(\dfrac{\partial \boldsymbol{H}_{\mathfrak{G},N_t}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}}\right)^{-1} \dfrac{\partial \boldsymbol{H}_{\mathfrak{G},N_t}}{\partial \boldsymbol{u}_{N_t}}\right]^{-\mathrm{T}} \\[4mm]
\qquad \cdot \left[\dfrac{\partial f_{int}}{\partial \boldsymbol{u}_{N_t}} - \bigwedge_{\text{el}} \sum_{\mathfrak{G}} \dfrac{\partial f_{int}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}} \left(\dfrac{\partial \boldsymbol{H}_{\mathfrak{G},N_t}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}}\right)^{-1} \dfrac{\partial \boldsymbol{H}_{\mathfrak{G},N_t}}{\partial \boldsymbol{u}_{N_t}}\right]^{\mathrm{T}} \\[4mm]
\boldsymbol{\gamma}_{\mathfrak{G},N_t} = -\left(\dfrac{\partial \boldsymbol{H}_{\mathfrak{G},N_t}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}}\right)^{-\mathrm{T}} \left(\dfrac{\partial f_{int}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}} + \boldsymbol{\lambda}_{N_t}^{\mathrm{T}} \dfrac{\partial \boldsymbol{R}_{N_t}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}}\right)^{\mathrm{T}} , \ \forall \ \mathfrak{G} \\[4mm]
\boldsymbol{\psi}_{N_t} = -\left(\dfrac{\partial \boldsymbol{G}_{N_t}}{\partial \boldsymbol{\omega}_{N_t}}\right)^{-\mathrm{T}} \left(\dfrac{\partial f_{int}}{\partial \boldsymbol{\omega}_{N_t}} + \boldsymbol{\lambda}_{N_t}^{\mathrm{T}} \dfrac{\partial \boldsymbol{R}_{N_t}}{\partial \boldsymbol{\omega}_{N_t}} + \bigwedge_{\text{el}} \sum_{\mathfrak{G}} \boldsymbol{\gamma}_{\mathfrak{G},N_t}^{\mathrm{T}} \dfrac{\partial \boldsymbol{H}_{\mathfrak{G},N_t}}{\partial \boldsymbol{\omega}_{N_t}}\right)^{\mathrm{T}}
\end{cases} \\[10mm]
n < N_t : \begin{cases}
\boldsymbol{\lambda}_n = -\left[\dfrac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{u}_n} - \bigwedge_{\text{el}} \sum_{\mathrm{G}} \dfrac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\nu}_{\mathrm{G},n}} \left(\dfrac{\partial \boldsymbol{H}_{\mathrm{G},n}}{\partial \boldsymbol{\nu}_{\mathrm{G},n}}\right)^{-1} \dfrac{\partial \boldsymbol{H}_{\mathrm{G},n}}{\partial \boldsymbol{u}_n}\right]^{-\mathrm{T}} \\[4mm]
\qquad \cdot \Bigg\{ \boldsymbol{\lambda}_{n+1}^{\mathrm{T}} \left[\dfrac{\partial \boldsymbol{R}_{n+1}}{\partial \boldsymbol{u}_n} - \bigwedge_{\text{el}} \sum_{\mathrm{G}} \dfrac{\partial \boldsymbol{R}_{n+1}}{\partial \boldsymbol{\nu}_{\mathrm{G},n}} \left(\dfrac{\partial \boldsymbol{H}_{\mathrm{G},n}}{\partial \boldsymbol{\nu}_{\mathrm{G},n}}\right)^{-1} \dfrac{\partial \boldsymbol{H}_{\mathrm{G},n}}{\partial \boldsymbol{u}_n}\right] \\[4mm]
\qquad \quad + \bigwedge_{\text{el}} \sum_{\mathrm{G}} \boldsymbol{\gamma}_{\mathrm{G},n+1}^{\mathrm{T}} \left[\dfrac{\partial \boldsymbol{H}_{\mathrm{G},n+1}}{\partial \boldsymbol{u}_n} - \dfrac{\partial \boldsymbol{H}_{n+1}}{\partial \boldsymbol{\nu}_{\mathrm{G},n}} \left(\dfrac{\partial \boldsymbol{H}_{\mathrm{G},n}}{\partial \boldsymbol{\nu}_{\mathrm{G},n}}\right)^{-1} \dfrac{\partial \boldsymbol{H}_{\mathrm{G},n}}{\partial \boldsymbol{u}_n}\right] \Bigg\}^{\mathrm{T}} \\[4mm]
\boldsymbol{\gamma}_{\mathrm{G},n} = -\left(\dfrac{\partial \boldsymbol{H}_{\mathrm{G},n}}{\partial \boldsymbol{\nu}_{\mathrm{G},n}}\right)^{-\mathrm{T}} \left(\boldsymbol{\lambda}_n^{\mathrm{T}} \dfrac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\nu}_{\mathrm{G},n}} + \boldsymbol{\lambda}_{n+1}^{\mathrm{T}} \dfrac{\partial \boldsymbol{R}_{n+1}}{\partial \boldsymbol{\nu}_{\mathrm{G},n}} + \boldsymbol{\gamma}_{\mathrm{G},n+1}^{\mathrm{T}} \dfrac{\partial \boldsymbol{H}_{\mathrm{G},n+1}}{\partial \boldsymbol{\nu}_{\mathrm{G},n}}\right)^{\mathrm{T}} , \ \forall \ \mathrm{G} \\[4mm]
\boldsymbol{\psi}_n = -\left(\dfrac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{\omega}_n}\right)^{-\mathrm{T}} \left[\boldsymbol{\lambda}_n^{\mathrm{T}} \dfrac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\omega}_n} + \bigwedge_{\text{el}} \sum_{\mathrm{G}} \left(\boldsymbol{\gamma}_n^{\mathrm{T}} \dfrac{\partial \boldsymbol{H}_{\mathrm{G},n}}{\partial \boldsymbol{\omega}_n} + \boldsymbol{\gamma}_{n+1}^{\mathrm{T}} \dfrac{\partial \boldsymbol{H}_{\mathrm{G},n+1}}{\partial \boldsymbol{\omega}_n}\right) + \boldsymbol{\psi}_{n+1}^{\mathrm{T}} \dfrac{\partial \boldsymbol{G}_{n+1}}{\partial \boldsymbol{\omega}_n}\right]^{\mathrm{T}}
\end{cases}
\end{aligned}
\tag{30}
$$

Since the adjoint vectors, $\boldsymbol{\lambda}_{N_t}$, $\boldsymbol{\gamma}_{N_t}$ and $\boldsymbol{\psi}_{N_t}$ at the end time step N_t are dependent on the tangent matrices of the residuals and the explicit derivatives $\frac{\partial f_{int}}{\partial \boldsymbol{u}_{N_t}}$, $\frac{\partial f_{int}}{\partial \boldsymbol{\nu}_{\mathrm{G},N_t}}$ and $\frac{\partial f_{int}}{\partial \boldsymbol{\omega}_{\mathrm{G},N_t}}$, we begin the sensitivity analysis by solving for the last adjoint vector in the time sequence starting from the final time step. For example, every other adjoint vector $\boldsymbol{\lambda}_n$ corresponding to the time step $n < N_t$, depends on the adjoint vector that follows it in the time sequence, $\boldsymbol{\lambda}_{n+1}$. Then the solved global mechanical adjoint vector $\boldsymbol{\lambda}_n$, together with $\boldsymbol{\gamma}_{n+1}$ and $\boldsymbol{\lambda}_{n+1}$ are used to solve for the local adjoint vectors $\boldsymbol{\gamma}_n$. Finally, the global thermal adjoint vector $\boldsymbol{\psi}_n$ is calculated with the solved global mechanical adjoint vector $\boldsymbol{\lambda}_n$ and local adjoint vector $\boldsymbol{\gamma}_n$, as well as information ($\boldsymbol{\lambda}_{n+1}$ and $\boldsymbol{\psi}_{n+1}$) from the later time step $n + 1$. Therefore, we solve for the remaining adjoint vectors in reverse chronological order until all adjoint responses are computed. Once all implicit terms have been eliminated using the above adjoint solution, the total sensitivity expression reduces to the

13

following.

$$
\begin{aligned}
\frac{\mathrm{d}\Pi}{\mathrm{d}\boldsymbol{r}} =& \frac{\partial f_{int}}{\partial \boldsymbol{r}} + \sum_{n=1}^{N_t} \left(\boldsymbol{\lambda}_n^{\mathrm{T}} \frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{r}} + \bigwedge_{\mathrm{el}} \sum_{\mathfrak{G}} \boldsymbol{\gamma}_{\mathfrak{G},n}^{\mathrm{T}} \frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{r}} + \boldsymbol{\psi}_n^{\mathrm{T}} \frac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{r}} \right) \\
& + \left[\boldsymbol{\lambda}_1^{\mathrm{T}} \left(\frac{\partial \boldsymbol{R}_1}{\partial \boldsymbol{u}_0} \frac{\mathrm{d}\boldsymbol{u}_0}{\mathrm{d}\boldsymbol{r}} + \bigwedge_{\mathrm{el}} \sum_{\mathfrak{G}} \frac{\partial \boldsymbol{R}_1}{\partial \boldsymbol{\nu}_{\mathfrak{G},0}} \frac{\mathrm{d}\boldsymbol{\nu}_{\mathfrak{G},0}}{\mathrm{d}\boldsymbol{r}} \right) \right. \\
& + \left. \bigwedge_{\mathrm{el}} \sum_{\mathfrak{G}} \boldsymbol{\gamma}_{\mathfrak{G},1}^{\mathrm{T}} \left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},1}}{\partial \boldsymbol{u}_0} \frac{\mathrm{d}\boldsymbol{u}_0}{\mathrm{d}\boldsymbol{r}} + \frac{\partial \boldsymbol{H}_{\mathfrak{G},1}}{\partial \boldsymbol{\nu}_{\mathfrak{G},0}} \frac{\mathrm{d}\boldsymbol{\nu}_{\mathfrak{G},0}}{\mathrm{d}\boldsymbol{r}} + \frac{\partial \boldsymbol{H}_{\mathfrak{G},1}}{\partial \boldsymbol{\omega}_0} \frac{\mathrm{d}\boldsymbol{\omega}_0}{\mathrm{d}\boldsymbol{r}} \right) + \boldsymbol{\psi}_1^{\mathrm{T}} \frac{\partial \boldsymbol{G}_1}{\partial \boldsymbol{\omega}_0} \frac{\mathrm{d}\boldsymbol{\omega}_0}{\mathrm{d}\boldsymbol{r}} \right]
\end{aligned}
\tag{31}
$$

Note that to solve for the sensitivity, the full history of global state variables (i.e. displacement field $\boldsymbol{d}$ and temperature field $\boldsymbol{T}$) and local state variables (i.e. stress tensor $\boldsymbol{\sigma}$, martensite volume fraction ξ, transformation strain tensor $\boldsymbol{\varepsilon}^t$, the compliance modulus S) are required. However, tangent matrices such as $\partial \boldsymbol{R}_n / \partial \boldsymbol{u}_n$, $\partial \boldsymbol{R}_n / \partial \boldsymbol{\nu}_n$ and etc. are only related to the current step, or one later step, or one future step; and can be iteratively stored and updated.

2.6.2. Analytical Solution of the Adjoint Vectors

For brevity, we present only the derivation of the adjoint parameters related to the active materials in this section. To compute the adjoint vectors $\boldsymbol{\lambda}_n$, $\boldsymbol{\gamma}_n$ and $\boldsymbol{\psi}_n$, thirteen tangent matrices, $\partial \boldsymbol{R}_n / \partial \boldsymbol{u}_n$, $\partial \boldsymbol{R}_n / \partial \boldsymbol{\nu}_n$, $\partial \boldsymbol{H}_n / \partial \boldsymbol{u}_n$, $\partial \boldsymbol{H}_n / \partial \boldsymbol{\nu}_n$, $\partial \boldsymbol{H}_n / \partial \boldsymbol{\omega}_n$, $\partial \boldsymbol{G}_n / \partial \boldsymbol{\omega}_n$, $\partial \boldsymbol{R}_{n+1} / \partial \boldsymbol{u}_n$, $\partial \boldsymbol{R}_{n+1} / \partial \boldsymbol{\nu}_n$, $\partial \boldsymbol{H}_{n+1} / \partial \boldsymbol{u}_n$, $\partial \boldsymbol{H}_{n+1} / \partial \boldsymbol{\nu}_n$, $\partial \boldsymbol{H}_{n+1} / \partial \boldsymbol{\omega}_n$ and $\partial \boldsymbol{G}_{n+1} / \partial \boldsymbol{\omega}_n$ in Equation 30 need to be evaluated. Note that the aforementioned matrices have different formulas depending on whether the SMAs behave elastically or inelastically. If the material exhibits inelastic behavior in the current time step, the local residual $\boldsymbol{H}$ should refer to the case where $\dot{\xi} \neq 0$ in Equation 18, and the form of the corresponding matrices should be as follows.

$$
\begin{aligned}
\frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{u}_n} =& \begin{bmatrix} \frac{\partial \boldsymbol{R}_n^c}{\partial \boldsymbol{F}_n^c} & \frac{\partial \boldsymbol{R}_n^c}{\partial \boldsymbol{d}_n^c} \\ \frac{\partial \boldsymbol{R}_n^f}{\partial \boldsymbol{F}_n^f} & \frac{\partial \boldsymbol{R}_n^f}{\partial \boldsymbol{d}_n^f} \end{bmatrix} = \begin{bmatrix} -\boldsymbol{I}^{cc} & \boldsymbol{0}^{cf} \\ \boldsymbol{0}^{fc} & \boldsymbol{0}^{ff} \end{bmatrix} \\
\frac{\partial \boldsymbol{R}_{\mathrm{el},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} =& \begin{bmatrix} \boldsymbol{0}_{24\times 8} & w\boldsymbol{B}^{\mathrm{T}}\det\boldsymbol{J} \end{bmatrix} \\
\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{u}_{\mathrm{el},n}} =& \begin{bmatrix} \boldsymbol{0}_{8\times c^{\mathrm{el}},} & \boldsymbol{0}_{8\times f^{\mathrm{el}}} \\ \boldsymbol{0}_{6\times c^{\mathrm{el}}} & \boldsymbol{S}_n^{-1} : \boldsymbol{B}_{f^{\mathrm{el}}} \end{bmatrix} \\
\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} =& \begin{bmatrix} \frac{\partial \boldsymbol{H}_\Phi^n}{\partial \xi_n} & \frac{\partial \boldsymbol{H}_\Phi^n}{\partial \boldsymbol{\varepsilon}^t} & \frac{\partial \boldsymbol{H}_\Phi^n}{\partial S_n} & \frac{\partial \boldsymbol{H}_\Phi^n}{\partial \boldsymbol{\sigma}_n} \\ \frac{\partial \boldsymbol{H}_{\boldsymbol{\varepsilon}^t}^n}{\partial \xi_n} & \frac{\partial \boldsymbol{H}_{\boldsymbol{\varepsilon}^t}^n}{\partial \boldsymbol{\varepsilon}^t} & \frac{\partial \boldsymbol{H}_{\boldsymbol{\varepsilon}^t}^n}{\partial S_n} & \frac{\partial \boldsymbol{H}_{\boldsymbol{\varepsilon}^t}^n}{\partial \boldsymbol{\sigma}_n} \\ \frac{\partial \boldsymbol{H}_\xi^n}{\partial \xi_n} & \frac{\partial \boldsymbol{H}_\xi^n}{\partial \boldsymbol{\varepsilon}^t} & \frac{\partial \boldsymbol{H}_\xi^n}{\partial S_n} & \frac{\partial \boldsymbol{H}_\xi^n}{\partial \boldsymbol{\sigma}_n} \\ \frac{\partial \boldsymbol{H}_{\boldsymbol{\sigma}}^n}{\partial \xi_n} & \frac{\partial \boldsymbol{H}_{\boldsymbol{\sigma}}^n}{\partial \boldsymbol{\varepsilon}_n^t} & \frac{\partial \boldsymbol{H}_{\boldsymbol{\sigma}}^n}{\partial S_n} & \frac{\partial \boldsymbol{H}_{\boldsymbol{\sigma}}^n}{\partial \boldsymbol{\sigma}_n} \end{bmatrix} \\
=& \begin{bmatrix} \partial_\xi \Phi_n & \boldsymbol{0}_{1\times 6} & 0 & \partial_{\boldsymbol{\sigma}} \Phi_n^{\mathrm{T}} \\ \boldsymbol{\Lambda}_n & -\boldsymbol{I}_{6\times 6} & \boldsymbol{0}_{6\times 1} & \partial_{\boldsymbol{\sigma}} \boldsymbol{\Lambda}_n : \Delta\xi_n \\ \Delta S & \boldsymbol{0}_{1\times 6} & -I_{1\times 1} & \boldsymbol{0}_{1\times 6} \\ \boldsymbol{0}_{6\times 1} & -\boldsymbol{S}_n^{-1} & -S_n^{-1} : \boldsymbol{\sigma}_n & -\boldsymbol{I}_{6\times 6} \end{bmatrix} \\
\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\omega}_{\mathrm{el},n}} =& \begin{bmatrix} \boldsymbol{0}_{1\times p^{\mathrm{el}}} & \partial_{\boldsymbol{T}} \Phi_n : \boldsymbol{N}_{d^{\mathrm{el}}} \\ \boldsymbol{0}_{7\times p^{\mathrm{el}}} & \boldsymbol{0}_{7\times d^{\mathrm{el}}} \\ \boldsymbol{0}_{6\times p^{\mathrm{el}}} & -\boldsymbol{S}_n^{-1} : \boldsymbol{\alpha} : \boldsymbol{N}_{d^{\mathrm{el}}} \end{bmatrix}
\end{aligned}
\tag{32}
$$

14

$$\frac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{\omega}_n} = \left[\begin{array}{cc} \frac{\partial \boldsymbol{G}_n^p}{\partial \boldsymbol{q}_n^p} & \frac{\partial \boldsymbol{G}_n^p}{\partial \boldsymbol{T}_n^p} \\ \frac{\partial \boldsymbol{G}_n^{in}}{\partial \boldsymbol{q}_n^{in}} & \frac{\partial \boldsymbol{G}_n^{in}}{\partial \boldsymbol{T}_n^{in}} \end{array} \right] = \left[\begin{array}{cc} -\boldsymbol{I}^{pp} & \widetilde{\boldsymbol{K}}_c^{\,pd} \\ \boldsymbol{0}^{dp} & \widetilde{\boldsymbol{K}}_c^{\,dd} \end{array} \right]$$

Again, the superscripts ^{f} and ^{c} refer to entries within the vectors that correspond to free and constrained degrees of freedom for the mechanical problem, while the superscripts ^{d} and ^{p} refer to entries within the vector that correspond to free and prescribed degrees of freedom for the thermal conduction problem, respectively. If the local material is still in an elastic state ($\dot{\xi} = 0$), the derivative of the residual for the current time step with respect to the current state variables will have a much simpler form.

$$\frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{u}_n} = \left[\begin{array}{cc} -\boldsymbol{I}^{cc} & \boldsymbol{0}^{cf} \\ \boldsymbol{0}^{fc} & \boldsymbol{0}^{ff} \end{array} \right]$$

$$\frac{\partial \boldsymbol{R}_{\mathrm{el},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} = \left[\begin{array}{c} w\boldsymbol{B}^{\mathrm{T}}\det\boldsymbol{J} \end{array} \right]$$

$$\frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{u}_{\mathrm{el},n}} = \left[\begin{array}{cc} \boldsymbol{0}_{6\times c^{\mathrm{el}}} & \boldsymbol{S}_n^{-1} : \boldsymbol{B}^{f\mathrm{el}} \end{array} \right] \tag{33}$$

$$\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} = \left[\begin{array}{c} -\boldsymbol{I}_{6\times 6} \end{array} \right]$$

$$\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\omega}_{\mathrm{el},n}} = \left[\begin{array}{cc} \boldsymbol{0}_{6\times p^{\mathrm{el}}} & -\boldsymbol{S}_n^{-1} : \boldsymbol{\alpha} : \boldsymbol{N}_{d^{\mathrm{el}}} \end{array} \right]$$

$$\frac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{\omega}_n} = \left[\begin{array}{cc} -\boldsymbol{I}^{pp} & \widetilde{\boldsymbol{K}}_c^{\,pd} \\ \boldsymbol{0}^{dp} & \widetilde{\boldsymbol{K}}_c^{\,dd} \end{array} \right]$$

Next, we discuss the tangent matrices at the $n+1$ time step, shown as below.

$$\frac{\partial \boldsymbol{R}_{n+1}}{\partial \boldsymbol{u}_n} = \boldsymbol{0}$$

$$\frac{\partial \boldsymbol{R}_{n+1}}{\partial \boldsymbol{\nu}_n} = \boldsymbol{0}$$

$$\frac{\partial \boldsymbol{H}_{\mathfrak{G},n+1}}{\partial \boldsymbol{u}_{\mathrm{el},n}} = \left\{ \begin{array}{ll} \left[\begin{array}{cc} \boldsymbol{0}_{8\times c^{\mathrm{el}}} & \boldsymbol{0}_{8\times f^{\mathrm{el}}} \\ \boldsymbol{0}_{6\times c^{\mathrm{el}}} & \boldsymbol{0}_{6\times f^{\mathrm{el}}} \end{array} \right] & \dot{\xi}_{n+1} \neq 0 \\ \left[\begin{array}{cc} \boldsymbol{0}_{14\times c^{\mathrm{el}}} & \boldsymbol{0}_{14\times f^{\mathrm{el}}} \end{array} \right] & \dot{\xi}_{n+1} = 0 \end{array} \right. \tag{34}$$

$$\frac{\partial \boldsymbol{H}_{\mathfrak{G},n+1}}{\partial \boldsymbol{\nu}_n} = \left\{ \begin{array}{ll} \left[\begin{array}{c} \boldsymbol{0}_{6\times 6} \end{array} \right] & \dot{\xi}_n = 0, \ \dot{\xi}_{n+1} = 0 \\ \left[\begin{array}{c} \boldsymbol{0}_{8\times 6} \\ \boldsymbol{0}_{6\times 6} \end{array} \right] & \dot{\xi}_n = 0, \ \dot{\xi}_{n+1} \neq 0 \\ \left[\begin{array}{cc} \boldsymbol{0}_{6\times 8} & \boldsymbol{0}_{6\times 6} \end{array} \right] & \dot{\xi}_n \neq 0, \ \dot{\xi}_{n+1} = 0 \\ \left[\begin{array}{cccc} 0 & \boldsymbol{0}_{1\times 6} & 0 & \boldsymbol{0}_{1\times 6} \\ -\boldsymbol{\Lambda}_{n+1} & \boldsymbol{I}_{6\times 6} & \boldsymbol{0}_{6\times 1} & \boldsymbol{0}_{6\times 6} \\ -\Delta S & \boldsymbol{0}_{1\times 6} & I_{1\times 1} & \boldsymbol{0}_{1\times 6} \\ \boldsymbol{0}_{6\times 1} & \boldsymbol{0}_{6\times 6} & \boldsymbol{0}_{6\times 1} & \boldsymbol{0}_{6\times 6} \end{array} \right] & \dot{\xi}_n \neq 0, \ \dot{\xi}_{n+1} \neq 0 \end{array} \right.$$

$$\frac{\partial \boldsymbol{H}_{\mathfrak{G},n+1}}{\partial \boldsymbol{\omega}_{\mathrm{el},n}} = \left\{ \begin{array}{ll} \left[\begin{array}{cc} \boldsymbol{0}_{14\times p^{\mathrm{el}}} & \boldsymbol{0}_{14\times d^{\mathrm{el}}} \end{array} \right] & \dot{\xi}_{n+1} \neq 0 \\ \left[\begin{array}{cc} \boldsymbol{0}_{6\times p^{\mathrm{el}}} & \boldsymbol{0}_{6\times d^{\mathrm{el}}} \end{array} \right] & \dot{\xi}_{n+1} = 0 \end{array} \right.$$

$$\frac{\partial \boldsymbol{G}_{n+1}}{\partial \boldsymbol{\omega}_n} = \left[\begin{array}{cc} \boldsymbol{0}^{pp} & -\widetilde{\boldsymbol{C}}^{pd} \\ \boldsymbol{0}^{dp} & -\widetilde{\boldsymbol{C}}^{dd} \end{array} \right]$$

15

Directly solving the algebraic equations for the adjoint vectors is computationally expensive. Moreover, computing the inverse of the matrix $\partial \boldsymbol{H}_n / \partial \boldsymbol{\nu}_n$ may result in ill-conditioned matrices, due to the large magnitude of the off-diagonal entries of the compliance matrix $\boldsymbol{S}$ and the compliance modulus S in $\partial \boldsymbol{H}_n / \partial \boldsymbol{\nu}_n$. Our previous work [38] has provided the analytical solution for evaluating the matrix, in the case where we adopt a flow rule that is independent of the stress tensor $\boldsymbol{\sigma}$. Herein, we discuss another case in which the flow rule is a function of stress and $\partial \boldsymbol{\Lambda}$ is nonzero, as shown in Equation 35. <mark>Note that when SMAs behave elastically, i.e. $\dot{\xi} = 0$, no transformation occurs and thus $\partial \boldsymbol{\Lambda}$ does not exist</mark> (Equation 11). Note also that $\boldsymbol{K} = \bigwedge_{\text{el}} \sum_{\text{G}} \boldsymbol{B}_{\text{G}}^{\text{T}} \boldsymbol{L}_n \boldsymbol{B}_{\text{G}}^{\text{T}} \det \boldsymbol{J}$ refers to the consistent tangent stiffness matrix in the mechanical problem. In addition, we provide a detailed derivation of the aforementioned analytical solution in the Appendix A. There we also show that the consistent tangent stiffness matrix $\boldsymbol{K}$ and the consistent tangent stiffness modulus $\mathfrak{L}_n$ in the sensitivity analysis are the same as those discussed in the finite element analysis [43].

$$\frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{u}_n} - \bigwedge_{\text{el}} \sum_{\mathfrak{G}} \frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\nu}_{\text{G},n}} \left(\frac{\partial \boldsymbol{H}_{\text{G},n}}{\partial \boldsymbol{\nu}_{\text{G},n}} \right)^{-1} \frac{\partial \boldsymbol{H}_{\text{G},n}}{\partial \boldsymbol{u}_{\text{el},n}} = \begin{bmatrix} -\boldsymbol{I}^{cc} & \boldsymbol{K}^{cf} \\ \boldsymbol{0}^{cf} & \boldsymbol{K}^{ff} \end{bmatrix}$$

$$\frac{\partial \boldsymbol{R}_{\text{el},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \right)^{-1} = w \boldsymbol{B}^{\text{T}} \begin{bmatrix} -\frac{\boldsymbol{\zeta}_n^{-1} : \partial_{\boldsymbol{\sigma}} \Phi_n}{Q} & \mathfrak{L}_n & \mathfrak{L}_n : \mathfrak{C}^{-1} : \boldsymbol{\sigma}_n & -\mathfrak{L}_n : \boldsymbol{S}_n \end{bmatrix} \det \boldsymbol{J}$$

$$\left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \right)^{-1} \frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{u}_{\text{el},n}} = \begin{bmatrix} \frac{\partial_{\boldsymbol{\sigma}} \Phi_n^{\text{T}} : \boldsymbol{\zeta}_n^{-1}}{Q} \\ \frac{\widetilde{\boldsymbol{\Lambda}}_n : \partial_{\boldsymbol{\sigma}} \Phi_n^{\text{T}}}{Q} : \boldsymbol{\zeta}_n^{-1} + \boldsymbol{S}_n : \boldsymbol{\zeta}_n^{-1} - \boldsymbol{I}_{6 \times 6} \\ \frac{\Delta S : \partial_{\boldsymbol{\sigma}} \Phi_n^{\text{T}}}{Q} : \boldsymbol{\zeta}_n^{-1} \\ -\mathfrak{L}_n \end{bmatrix} \boldsymbol{B} \qquad (35)$$

$$\frac{\partial \boldsymbol{H}_{\mathfrak{G},n+1}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \right)^{-1} = \begin{bmatrix} \mathfrak{A}_{8 \times 8} & \mathfrak{B}_{8 \times 6} \\ \boldsymbol{0}_{6 \times 8} & \boldsymbol{0}_{6 \times 6} \end{bmatrix}$$

$$\frac{\partial \boldsymbol{H}_{\mathfrak{G},n+1}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \right)^{-1} \frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{u}_{\text{el},n}} = \begin{bmatrix} \boldsymbol{0}_{1 \times 6} \\ \frac{\widetilde{\boldsymbol{\Lambda}}_n - \boldsymbol{\Lambda}_{n+1}}{Q} : \partial_{\boldsymbol{\sigma}} \Phi_n^{\text{T}} : \boldsymbol{\zeta}_n^{-1} + \boldsymbol{S}_n : \boldsymbol{\zeta}_n^{-1} - \boldsymbol{I}_{6 \times 6} \\ \boldsymbol{0}_{7 \times 6} \end{bmatrix} \boldsymbol{B}$$

where

$$\mathfrak{A} = \begin{bmatrix} 0_{1 \times 1} & \boldsymbol{0}_{1 \times 7} \\ \frac{\widetilde{\boldsymbol{\Lambda}}_n - \boldsymbol{\Lambda}_{n+1}}{Q} & \\ 0_{1 \times 1} & \mathfrak{a}_{7 \times 7} \end{bmatrix}$$

$$\mathfrak{a} = \begin{bmatrix} (-\boldsymbol{S}_n - \frac{\widetilde{\boldsymbol{\Lambda}}_n - \boldsymbol{\Lambda}_{n+1}}{Q} : \partial_{\boldsymbol{\sigma}} \Phi_n^{\text{T}}) : \boldsymbol{\zeta}_n^{-1} & (-\frac{\widetilde{\boldsymbol{\Lambda}}_n - \boldsymbol{\Lambda}_{n+1}}{Q} : \partial_{\boldsymbol{\sigma}} \Phi_n^{\text{T}} : \boldsymbol{\zeta}_n^{-1} : \boldsymbol{S}_n + \boldsymbol{S}_n : \boldsymbol{\zeta}_n^{-1} : \partial_{\boldsymbol{\sigma}} \boldsymbol{\Lambda}_n : \Delta \xi_n) : S_n^{-1} : \boldsymbol{\sigma}_n \\ \boldsymbol{0}_{1 \times 6} & -I_{1 \times 1} \end{bmatrix}$$

$$\mathfrak{B} = \begin{bmatrix} \boldsymbol{0}_{1 \times 6} \\ \frac{\widetilde{\boldsymbol{\Lambda}}_n - \boldsymbol{\Lambda}_{n+1}}{Q} : \partial_{\boldsymbol{\sigma}} \Phi_n^{\text{T}} : \boldsymbol{\zeta}_n^{-1} + \boldsymbol{S}_n : \boldsymbol{\zeta}_n^{-1} - \boldsymbol{I}_{6 \times 6} \\ \boldsymbol{0}_{1 \times 6} \end{bmatrix} : \boldsymbol{S}_n \qquad (36)$$

$$\widetilde{\boldsymbol{\Lambda}}_n = \boldsymbol{S}_n : \boldsymbol{\zeta}_n^{-1} : (\boldsymbol{\Lambda}_n - \partial_{\boldsymbol{\sigma}} \boldsymbol{\Lambda}_n : \Delta \xi_n : S_n^{-1} : \boldsymbol{\sigma}_n : \Delta S)$$

In Table 1, we provide a pseudo-code of the algorithmic implementation for evaluating the sensitivities of SMA-based structure considering thermomechanical coupling effects.

I. Calculate adjoint vectors

16

Case A. $n = N_t$

  1. Let $n = N_t$

  2. Mechanical Global Level

    1) Let $-\bigwedge\limits_{\text{el}} \sum\limits_{\mathfrak{G}} \frac{\partial \boldsymbol{R}_{N_t}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}} \left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},N_t}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}} \right)^{-1} \frac{\partial \boldsymbol{H}_{\mathfrak{G},N_t}}{\partial \boldsymbol{u}_{\text{el},N_t}} = \boldsymbol{K}_{N_t}$, calculate $\frac{\partial f_{int}}{\partial \boldsymbol{u}_{N_t}}$

    2) Loop over elements, calculate $\frac{\partial f_{int}}{\partial \boldsymbol{\nu}_{N_t}} \left(\frac{\partial \boldsymbol{H}_{N_t}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}} \right)^{-1} \frac{\partial \boldsymbol{H}_{N_t}}{\partial \boldsymbol{u}_{\text{el},N_t}}$, $\frac{\partial f_{int}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}} \left(\frac{\partial \boldsymbol{H}_{N_t}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}} \right)^{-1}$ and

      $\frac{\partial \boldsymbol{R}_{N_t}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}} \left(\frac{\partial \boldsymbol{H}_{N_t}}{\partial \boldsymbol{\nu}_{\mathfrak{G},N_t}} \right)^{-1}$

    3) Assemble to get global $\frac{\partial f_{int}}{\partial \boldsymbol{\nu}_{N_t}} \left(\frac{\partial \boldsymbol{H}_{N_t}}{\partial \boldsymbol{\nu}_{N_t}} \right)^{-1} \frac{\partial \boldsymbol{H}_{N_t}}{\partial \boldsymbol{u}_{N_t}}$

    4) Calculate $\boldsymbol{\lambda}_{N_t}$

  3. Mechanical Local Level

    1) Loop over elements, calculate $\boldsymbol{\gamma}_{\mathfrak{G},N_t}$ for each Gauss point

  4. Thermal Global Level

    1) Calculate $\frac{\partial \boldsymbol{G}_{N_t}}{\partial \boldsymbol{\omega}_{N_t}} = \frac{\partial \boldsymbol{G}_n}{\partial \boldsymbol{\omega}_n}$, $\frac{\partial \boldsymbol{G}_{n+1}}{\partial \boldsymbol{\omega}_n}$ and $\frac{\partial f_{int}}{\partial \boldsymbol{\omega}_{N_t}}$

    2) Loop over elements, calculate $\boldsymbol{\gamma}_{\mathfrak{G},N_t}^{\text{T}} \frac{\partial \boldsymbol{H}_{\mathfrak{G},N_t}}{\partial \boldsymbol{\omega}_{N_t}}$ at each Gauss point and $\boldsymbol{\lambda}_{\text{el}}^{\text{T}} \frac{\partial \boldsymbol{R}_{N_t}}{\partial \boldsymbol{\omega}_{\text{el},N_t}}$

      for each element

    3) Assemble to get global $\bigwedge\limits_{\text{el}} \sum\limits_{\mathfrak{G}} \boldsymbol{\gamma}_{\mathfrak{G},N_t}^{\text{T}} \frac{\partial \boldsymbol{H}_{\mathfrak{G},N_t}}{\partial \boldsymbol{\omega}_{N_t}}$ and $\boldsymbol{\lambda}^{\text{T}} \frac{\partial \boldsymbol{R}_{N_t}}{\partial \boldsymbol{\omega}_{N_t}}$, calculate $\boldsymbol{\psi}_{N_t}$

    4) Let $n = n - 1$, go to Case B.

Case B. $n < N_t$

  1. Mechanical Global Level

    1) Let $-\bigwedge\limits_{\text{el}} \sum\limits_{\mathfrak{G}} \frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \right)^{-1} \frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{u}_{\text{el},n}} = \boldsymbol{K}_n$, calculate $\frac{\partial \boldsymbol{R}_{n+1}}{\partial \boldsymbol{u}_n}$

    2) Loop over elements, calculate $\frac{\partial \boldsymbol{R}_{n+1}}{\partial \boldsymbol{\nu}_n} \left(\frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{\nu}_n} \right)^{-1} \frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{u}_n}$, $\frac{\partial \boldsymbol{H}_{n+1}}{\partial \boldsymbol{u}_n}$, $\frac{\partial \boldsymbol{H}_{n+1}}{\partial \boldsymbol{\nu}_n} \left(\frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{\nu}_n} \right)^{-1} \frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{u}_n}$

      $\frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\nu}_n} \left(\frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{\nu}_n} \right)^{-1}$, $\frac{\partial \boldsymbol{R}_{n+1}}{\partial \boldsymbol{\nu}_n} \left(\frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{\nu}_n} \right)^{-1}$, $\frac{\partial \boldsymbol{H}_{n+1}}{\partial \boldsymbol{\nu}_n} \left(\frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{\nu}_n} \right)^{-1}$

    3) Assemble $\frac{\partial \boldsymbol{R}_{n+1}}{\partial \boldsymbol{\nu}_n} \left(\frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{\nu}_n} \right)^{-1} \frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{u}_n}$, $\frac{\partial \boldsymbol{H}_{n+1}}{\partial \boldsymbol{u}_n}$, $\frac{\partial \boldsymbol{H}_{n+1}}{\partial \boldsymbol{\nu}_n} \left(\frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{\nu}_n} \right)^{-1} \frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{u}_n}$

    4) Calculate $\boldsymbol{\lambda}_n$

  2. Mechanical Local Level

    1) Loop over elements, calculate $\boldsymbol{\gamma}_{\mathfrak{G},n}$ for each Gauss point

  3. Thermal Global Level

    1) Calculate $\boldsymbol{\psi}_{n+1}^{\text{T}} \frac{\partial \boldsymbol{G}_{n+1}}{\partial \boldsymbol{\omega}_n}$

    2) Loop over elements, calculate $\boldsymbol{\gamma}_{\mathfrak{G},n}^{\text{T}} \frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\omega}_n}$, $\boldsymbol{\gamma}_{\mathfrak{G},n+1}^{\text{T}} \frac{\partial \boldsymbol{H}_{\mathfrak{G},n+1}}{\partial \boldsymbol{\omega}_n}$ at each Gauss point, and

      $\boldsymbol{\lambda}_{\text{el}}^{\text{T}} \frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\omega}_{\text{el},n}}$ for each element

    3) Assemble to get global $\bigwedge\limits_{\text{el}} \sum\limits_{\mathfrak{G}} (\boldsymbol{\gamma}_{\mathfrak{G},n}^{\text{T}} \frac{\partial \boldsymbol{H}_{\text{G},n}}{\partial \boldsymbol{\omega}_n} + \boldsymbol{\gamma}_{\mathfrak{G},n+1}^{\text{T}} \frac{\partial \boldsymbol{H}_{\mathfrak{G},n+1}}{\partial \boldsymbol{\omega}_n})$ and $\boldsymbol{\lambda}^{\text{T}} \frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\omega}_n}$, calculate $\boldsymbol{\psi}_n$

    4) Check termination condition

      If $n > 0$

        Let $n = n - 1$, go back to Case B.

      else

        go to step II.

II. Calculate sensitivity

Calculate sensitivities according to equation 31.

Table 1: Implementation procedure for adjoint sensitivity analysis of SMAs

17

Based on our numerical tests, 2D finite element analysis of a bi-material mesh, which contains SMAs and an inactive material with 1000 quadrilateral elements, takes approximately 2 seconds to converge for one time step. These tests were implemented via the Matlab software on a platform that uses an Intel i5 CPU and contains 16GB of ROM. For the adjoint sensitivity analysis, it takes approximately one-third of the time required by the finite element analysis. Note that the sensitivity analysis requires storing the full history of state and adjoint variables. This feature makes the algorithm memory-intensive. However, calculating the adjoint vectors at time step n only requires the tangent of matrices from the current and subsequent time steps (i.e. $\partial \boldsymbol{H}_{n+1}/\partial \boldsymbol{\nu}_n$ and $\partial \boldsymbol{H}_n/\partial \boldsymbol{\nu}_n$). Therefore, to save memory, this information can be updated only as needed without storing the full history. In practice, we generally choose time steps that lead to a $0.1K - 0.5K$ temperature increment in the simulation. Based on our numerical tests, such time steps provide a suitable balance between the stability of the algorithm and the memory requirements for the storage of the internal state variables. Note that reducing the time step size by half could lead to a ten-fold increase in the storage requirement, as the full history of the internal state variables and the adjoint vectors are required to be stored.

Figure 6 shows the relative error $e = |ANA - FD|/|FD|$ between the sensitivity calculated from a finite difference approach (i.e. FD) and the analytical approach proposed in this work (i.e. ANA). The results indicate a good match between the two approaches. In addition, we can also observe that our sensitivity formulations enable the same level of precision when evaluating derivatives of inelastic behaviors of SMAs (i.e. $0 < \xi < 1$) evaluating derivatives of linear thermoelastic behaviors (i.e. $\xi = 0$ or $\xi = 1$). Details of the sensitivity data are presented in Appendix B from Table B.4 to Table B.9.

(a) Simulating TWSME
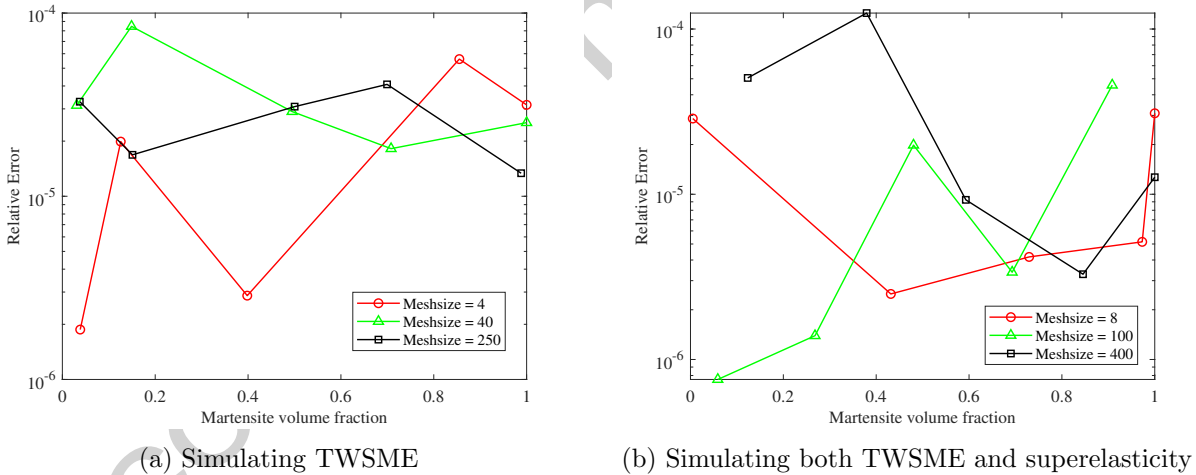
(b) Simulating both TWSME and superelasticity

Figure 6: Relative error between sensitivity calculated by finite differences and analytical solutions

For the optimization algorithm, we choose the standard method of moving asymptotes (MMA) [47]. This method allows for solving convex approximations of the nonlinear optimization problems.

3. Example Problems

Herein, we use the quadratic hardening function [5] to model the hardening process due to phase transformation of SMAs,

$$\text{Quadratic}: \ f(\xi) = \begin{cases} \frac{1}{2}\rho b^M \xi^2 + (\mu_1 + \mu_2)\xi \ (\dot{\xi} > 0) \\ \frac{1}{2}\rho b^A \xi^2 + (\mu_1 - \mu_2)\xi \ (\dot{\xi} < 0) \end{cases} \tag{37}$$

18

	Austenite (A)	Martensite (M)
Young's modulus E (Pa)	32.5×10^9	23.0×10^9
Thermal expansion coefficient α (K^{-1})	22.0×10^{-6}	22.0×10^{-6}
Specific heat c (J/kgK)	400.0	400.0
Transformation start temperature (K)	241	226
Transformation end temperature (K)	290	194
Highest transformation strain H	0.033	
Material density ρ (kg/m^3)	6500	
Reference temperature T_0 (K)	300	
Entropy difference $\rho\Delta s_0$ (J/m^{3}K)	-11.55×10^4	
Thermal conductivity κ (W/(m·K))	18	
Melting point (K)	$1523 \sim 1553$	

Table 2: Material properties of NiTi50

where b^M, b^A, μ_1 and μ_2, are constants that depend on the intrinsic material properties [5]. Table 2 contains the material properties of nickel titanium (NiTi50) shape-memory alloy that we use for simulation throughout the paper.

Figure 7a shows 1D TWSME of the quadratic hardening process under a uniform temperature cycle, i.e. variation of transformation strain (ε^t) and martensite volume fraction (ξ) in terms of temperature. Figure 7b presents the key evolution relationships of superelasticity under a uniformly distributed stress cycle, i.e. variation of transformation stress and martensite volume fraction in terms of stress. However, nonuniform temperature field and nonuniform stress field, which are common in real applications, will make the hardening process, i.e. the curves in Figure 7, more nonlinear. The nonuniform environmental conditions will gradually change the transformation temperature and stress during the transformation process (see Figure 3). This phenomenon is common throughout this paper and will be discussed in later results.

Note also that the martensite volume fraction and transformation strain are monotonic functions throughout the transformation process, which can be observed from Figure 7. Therefore, in the process of topology optimization, we suggest simulating the structure for a short period of time steps when most of SMA material begins its transformation (i.e. inelastic behaviors dominate the thermomechanical responses of the structure) to increase computational efficiency. Another reason for the suggestion lies in the fact that nonuniform thermomechanical conditions will change the start and end points of the transformation temperature and transformation stress ranges, as we have discussed in Figure 3. As a result, it would be difficult to assume the exact transformation end temperature for a local integration point (i.e. Gauss point), based on the material properties. Hence, it is necessary to simulate over a duration of time during which the transformation is most likely to occur. In addition, we apply a constant external load for the design that ensures a requisite internal stress. This step is necessary for the flow rule of the forward transformation used in this paper (see Equation 11), which requires a nonzero stress condition to activate the transformation process. Nevertheless, the proposed framework is generic and applicable to other flow rules that may work with zero-stress conditions.

Table 3 presents the material properties of the inactive material candidates used in this paper. For physical properties having a range of values, we adopt its lower bound in designing topology optimization problems. As noted earlier, we assume that all the inactive materials have the same thermal expansion coefficient(α), specific heat (c) and Poisson' ratio v as the NiTi50 for simplicity. We also assume SAC305 and BC have the same material density as NiTi50, as the relative difference

19

(a) Two-way shape-memory effect
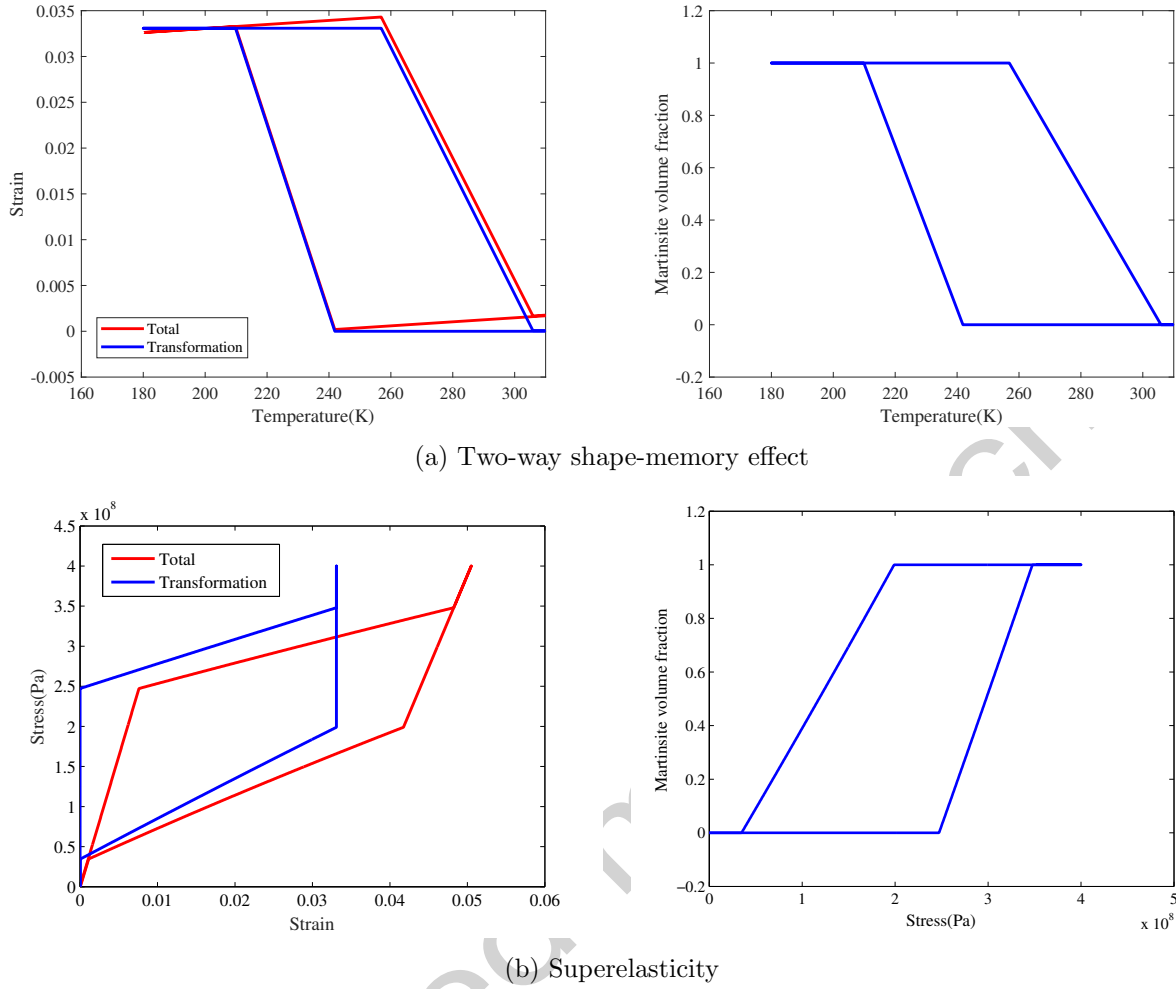

(b) Superelasticity

Figure 7: Simulation of the TWSME and pseudoelasticity in NiTi50 [5, 48]

of material density of SAC305 and BC with NiTi50 are within 5% and 15%, respectively.

Material	Void	SAC305	BC
Young's modulus E (Pa)	1	51×10^9	131×10^9
Thermal conductivity κ (W/(m·K))	0.5	59	105
Material density ρ (kg/m^3)	1.15	6500	6500
Melting point (K)	-	$490 \sim 491$	$1138 \sim 1228$

Table 3: Material properties of inactive materials [48]

3.1. Bi-material Active Bending Structure

In this section, we present the design of an active bending cantilevered beam structure composed of SMA and an inactive material BC. Properties of the two materials are stated in Table 2 and 3.

20

The optimization problem is stated in Equation 38

$$\begin{aligned} \min \quad & f_{obj} = -|\boldsymbol{L}^{\mathrm{T}}\boldsymbol{d}|, \\ s.t \quad & V_{\mathrm{SMA}} \le 0.3 \end{aligned} \tag{38}$$

where we want to maximize the vertical displacement at point A triggered by the TWSME, while constraining the use of SMAs to be 30% of the volume of whole structure. The mechanical and thermal boundary conditions of the design domain are shown in Figure 8.

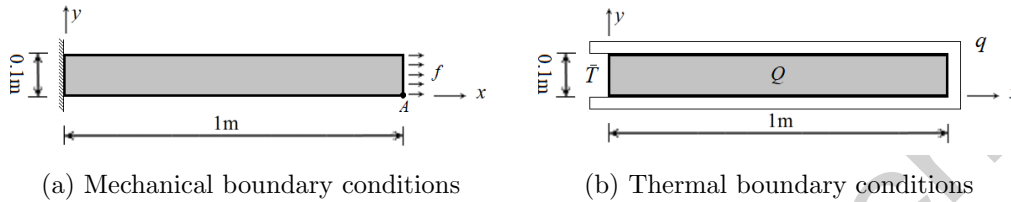(a) Mechanical boundary conditions   (b) Thermal boundary conditions

Figure 8: Boundary conditions for topology optimization of an active bending beam

For the mechanical boundary condition, a constant uniform surface traction $\boldsymbol{f} = 9 \times 10^7 \mathrm{N/m}$ is applied when cooling the beam. Simulation tests show that the transformation start temperature of the martensite phase $(M_s^{\boldsymbol{\sigma}})$ is increased to 250K under the applied constant load. Note that $M_s^{\boldsymbol{\sigma}}$ can be different for SMAs at different locations of the structure due to the nonuniform stress and the nonuniform temperature field. Thus, we set the possible highest transformation start temperature $250K$ as the initial and the reference temperature field of the beam. The cooling process is discretized into 25 time steps. The imposed heat flux is set to increase incrementally to from zero to $q = 2.5 \times 10^3$ w/m. In addition, we apply a heat sink throughout the domain of the structure, which represents the convective heat exchange between the structure's 2D surface in the xy-plane with the outside environment. The heat sink is set to decrease to from zero to $Q = -1.25 \times 10^5$ w/m^2 during the cooling process. The temperature $\bar{T}$ at the left side of the bar is held at 250K during the cooling process. The volumetric heat generation and heat flux are intended to mimic the process of immersing the structure into cold liquid or air. Water or air cooling and heating are a common approaches to activate SMAs and are suitable for building our framework. The process through which heat is conducted to the SMA structure through its surface in the xy-plane is mathematically similar to 2D volumetric heat generation.

The mesh used for simulation contains 4000 quadrilateral elements. Each element's effective Young's modulus, thermal conductivity and material density are calculated via Equation 22 and 23. The penalization factor P is chosen to be 3, and we implement a linear density filter with a radius equivalent to twice of the minimum distance between the centers of two neighbor elements. The material distribution of the optimization results, in terms of Young's modulus, is shown in Figure 9, where the blue region represents the SMA material and the red region represents the inactive material. B, C, D are three points of interest in observing the variation of inelastic behaviors in the structure. In this result, we observe a sawtooth-like pattern at the material interface. It is inconclusive whether this pattern is a mathematical artifact (such as a local minimum) or whether it confers improved mechanical performance. However, through our numerical tests, we notice that if the stiffness of the two material candidates are relatively close in value (i.e. NiTi50 and Ti64), the sawtooth-like layouts will become smoother, suggesting that this pattern enables the structure to leverage the differences in material properties.

Figure 10 shows the corresponding convergence plot of the topology optimization.

Figure 11 presents the topology optimization results using coarse meshes. Compared with the fine-mesh design (i.e. Figure 9), we observe that the general material layouts are the same despite
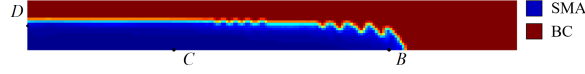
21

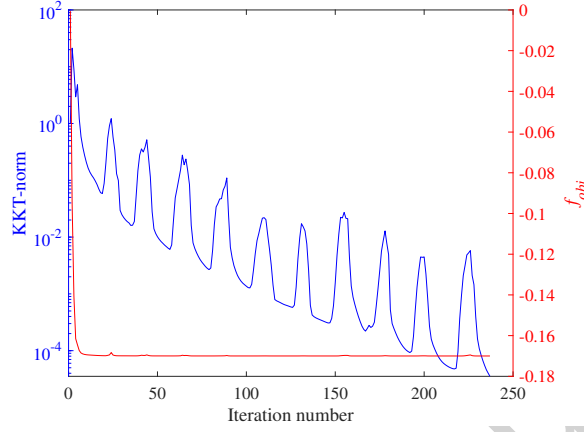Figure 9: Material distribution of the optimal bi-material bending beam

Figure 10: Convergence history during the optimization of the bi-material bending beam

the different mesh sizes. This result indicates that density filter is functioning as intended and the solution is mesh-independent.

(a) Mesh size = 250   (b) Mesh size = 1000

Figure 11: Mesh-dependency of the bending design

After obtaining the optimal design, we cool the structure for 200 time steps to observe long-term behaviors of the structure caused by the TWSME. Figure 12 shows the deformation, the corresponding temperature field, and the tip displace placement in the vertical direction at two sample time steps. Figure 13 gives details of the displacement history and quantifies the respective contributions from the thermoelastic and inelastic behaviors. In Figure 13, displacement due to different thermomechanical behaviors are calculated in accordance with the strain relationship shown in Equation 9. "Total", "Thermoelastic" and "Transformation" refer to the displacement related to the total strain (ε), the displacement related to the thermoelastic strain ($\varepsilon^e + \varepsilon^{th}$) and the displacement related to the transformation strain (ε^t), respectively.

From Figure 13, we observe a plateau of the tip displacement in the vertical direction during the first 20 time steps. This result indicates that the structural deformation is dominated by the thermoelastic behaviours of the beam at the early cooling stage, thus the beam only has minimum bending deformation due to thermoelastic expansion. As the beam is gradually cooled down, the transformation displacement dominates the tip displacement. Hence, we notice much larger bending of the beam caused by the TWSME, which cannot be achieved by pure thermoelastic deformation.

Figure 14 shows the nodal temperature, Gauss-point martensite volume fraction (ξ), Gaussian-point horizontal transformation strain (ε^t_x) and Gaussian-point Young's modulus of SMA ($1/S$) varied in time. We evaluate these relationships at mesh nodes and Gauss points that are closest to
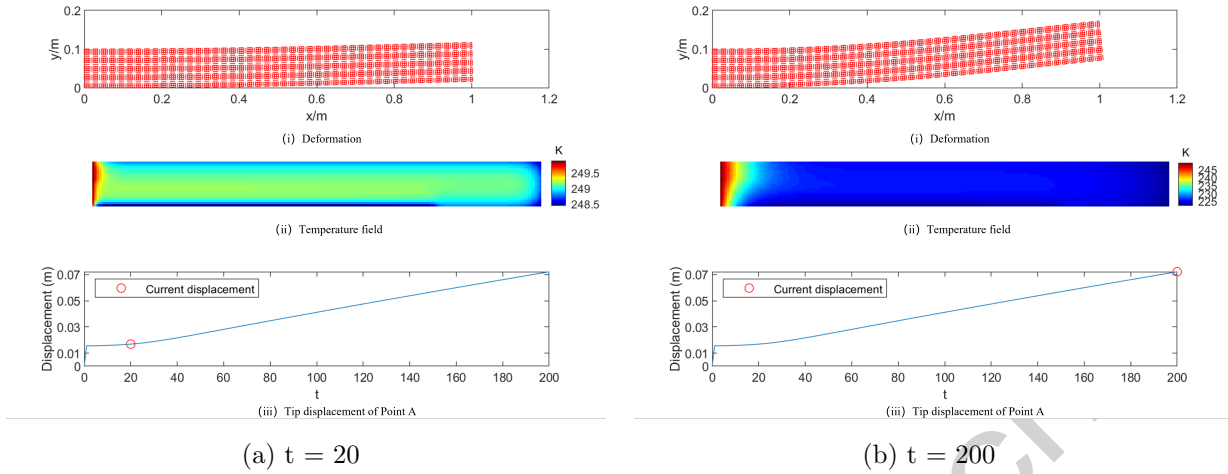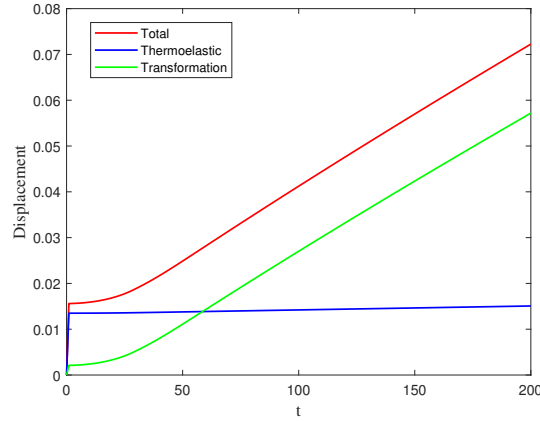
22

Figure 12: Deformation of the bimaterial beam during the cooling process

Figure 13: Contribution of different thermomechanical behaviors to the tip displacement of point A in the y-direction

points B, C and D (cf. Figure 9). Evolution of the martensite volume fraction, vertical transformation strain and Young's modulus of SMAs reveal that the forward transformation of the three points happens sequentially. Hence the quadratic hardening process becomes more nonlinear than the benchmark results illustrated in Figure 7. In addition, we observe that an inefficient cooling process with a minimum temperature change will not trigger the transformation (i.e. point C).

In short, the optimal design of the bi-material beam successfully leverages the inelastic behaviors of SMAs. The optimization algorithm captures the sequential transformation process at different locations, to maximize the overall inelastic performance of the structure. As a result, we obtain an optimal design that shows favorable active bending behaviors mainly due to the TWSME.

3.2. Multi-material Gripper Design

In this section, we present the design of a self-actuating SMA-based grippers having both superelasticity and TWSME. The gripper is composed of multiple material candidates, including SMA, SAC305 and BC, plus a void material candidate. Properties of the three material candidates

23

(a) Temperature

(b) Martensite volume fraction relation

(c) Transformation strain ε_x^t relation
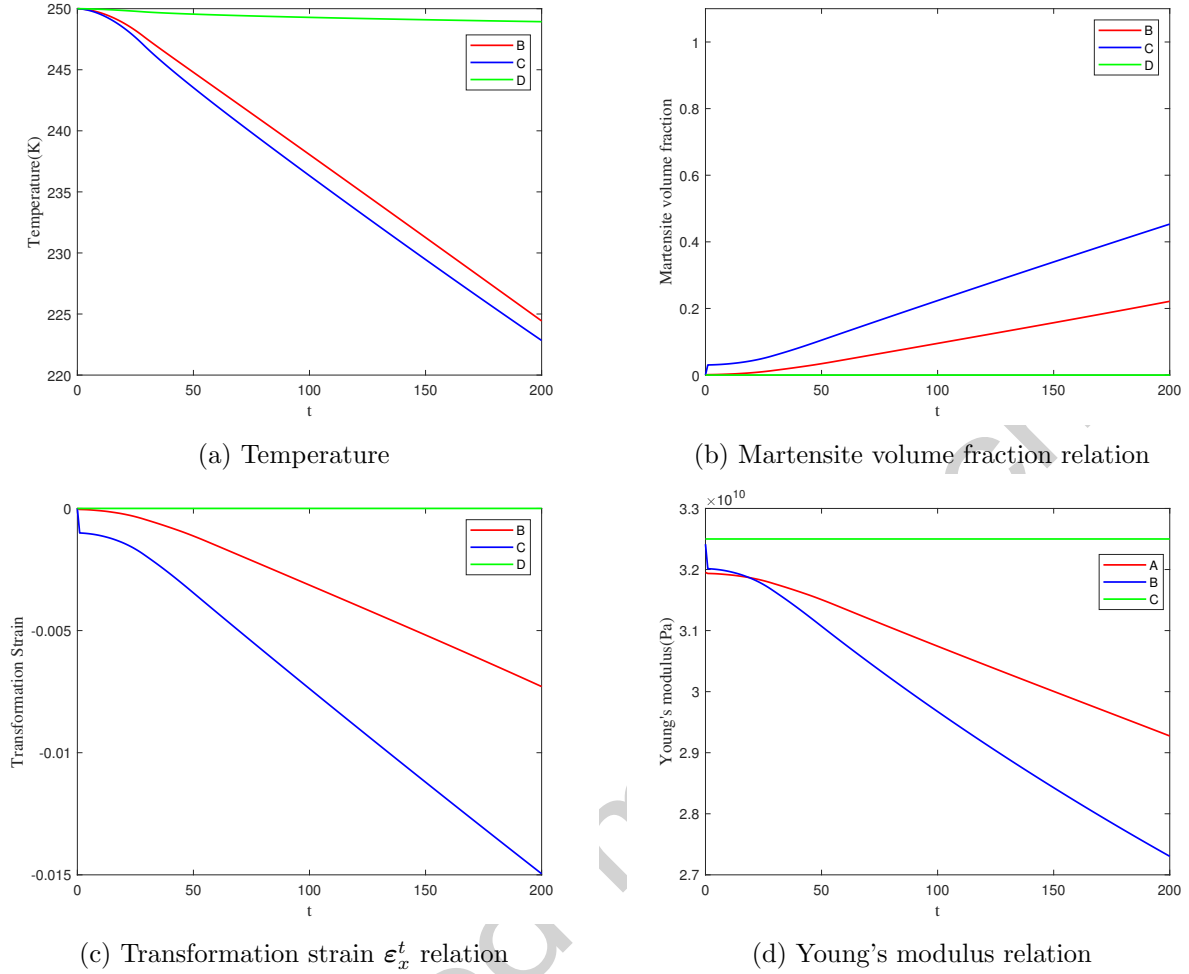
(d) Young's modulus relation

Figure 14: Variation of state variables as a function of time computed at three points of interest within the structure

(plus void) are stated in Table 2 and 3. The optimization problem is stated in Equation 39

$$
\begin{aligned}
\min \quad & f_{obj} = -|k_s d_y^A|, \\
s.t \quad & d_{in} \leq 8 \times 10^{-4}\text{m} \\
& g = \left[\sum_{i=1}^{Ne} \left(\sum_{j=2}^{4} (\mu_i^{(j)})^\eta \frac{T_{N_t}^{(i)}}{T^{*(j)}} \right)^\beta \right]^{1/\beta} \leq 1 \\
& M = \frac{\sum_{i=1}^{Ne} \sum_{j=2}^{4} \mu_i^{(j)} E^{(j)}}{E_{max} \cdot Ne} \leq 0.15
\end{aligned}
\tag{39}
$$

where k_s is the spring stiffness of the output spring and we want to maximize the vertical output force of point A triggered by the TWSME and superelasticity (note that we assume symmetry so we design and simulate only the top half of the mechanism). Here d_{in} is the initial (instantaneous) displacement, which is constrained to 8×10^{-4}m. This constraint is used to avoid an overly compliant design. An overly compliant structure is more susceptible to large deformations and may cause the finite element model to diverge. It may also lead to a non-convergent design,

24

which is characterized by the presence of intermediate materials and blurry material interfaces. Hence, a stiffness constraint is used to mitigate these issues. The function g is a material-specific temperature constraint used to prevent melting or viscoelastic creep within the material under the applied thermal environment, where $T^{*(j)}$ is a material-specific temperature limit for each design material (excluding void). In addition, we defined a constraint representing the relative mass of the structure M. Here the function M quantifies the amount of non-void material in the structure as a fraction of to maximum possible material usage. Within this equation, material with larger stiffness (represented by E), is penalized with increased relative mass. In this problem, the relative mass limit is set to be 0.15 to save the usage of solid materials especially SMAs to save cost. The mechanical and thermal boundary conditions of the design domain are shown in Figure 15.
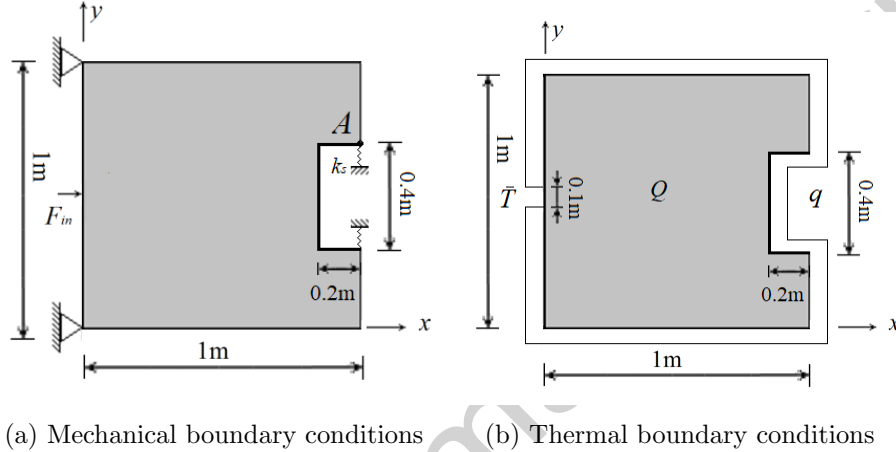

(a) Mechanical boundary conditions   (b) Thermal boundary conditions

Figure 15: Boundary conditions for topology optimization of a multimaterial gripper

On the left side of the design domain, a input force $\boldsymbol{F}_{in}$ increasing from zero to 1.25×10^8N is applied to guarantee the activation of both TWSME and superelasticity. Two springs, with stiffness $k_s = 5$N/m, are attached to the tips of the gripper. The structure is subjected to an initial temperature field of $230K$, with a reference temperature of $240K$. To trigger the TWSME, a cooling process over 20 time steps is implemented. The outward heat flux of the cooling process is set to increase incrementally to $q = 2.5 \times 10^3$ w/m, with the heat sink decreasing to a maximum of $Q = -1.25 \times 10^5$ w/m^2. The temperature $\bar{T}$ at the left side of the domain remains at 230K during the cooling process. In addition, the parameter β in the thermal performance constraint g, c.f. Equation 40, is chosen to be 10. This parameter is used to obtain a smooth approximation of the maximum elemental temperature. The material-specific temperature constraint T^* is set to be 17K for SAC305, 227.6K for SMA, and 304K for BC, which are 10% (SAC305) and 20% (SMA, BC) of each material's melting point, cf. Table 2 and 3. The set of material-specific constraints are based on precept that material creeping point is generally 10 to 30% of the melting point of the materials (measured in the Kelvin scale) [48]. Note that we choose a lower threshold 10% of the melting point for the creep constraint of SAC305, since SAC305 has a tendency toward brittle fracture due to thermal shock [49]. The penalization factor η in the thermal performance constraint g is selected to be 0.8 to relax the intermediate density at the boundary of the non-void design domain. More details about sensitivity analysis of the thermal constraint g can be found in our previous work [46].

During the topology optimization, we use a mesh of 5888 quadrilateral elements for finite element analysis. The effective Young's modulus and thermal conductivity are calculated via Equation 22 and 24. The penalization factor P is chosen to be 3, and the radius of the density filter is twice of

25

the minimum distance between the centers of two neighbor elements. The optimization results are shown in Figure 16, where again B, C, D are three points of interest in sampling the inelastic behaviors of SMAs across the structure. Note that though we have three material candidates plus void, the optimizer only chooses SMA, BC and void to construct the design. The result demonstrates the effectiveness of the material-specific thermal performance constraint. Since the temperature constraint of SAC305 (17K) is much lower than the possible transformation temperature of SMAs, SAC305 is omitted from the design. Note that the material distribution is mainly determined by the material-specific temperature constraint and possible stress concentration, and also influenced by the thermal boundary conditions. Figure 17 shows the corresponding convergence plots of the topology optimization.
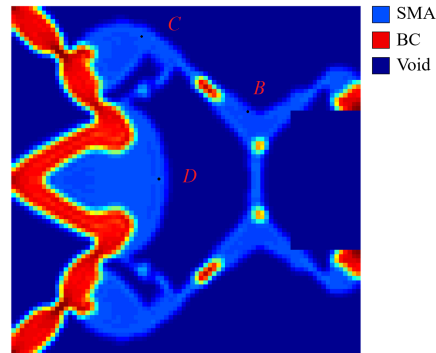
Figure 16: Optimized material distribution of a multi-material gripper

(a) Convergence history of the objective

(b) Convergence history of the constraints
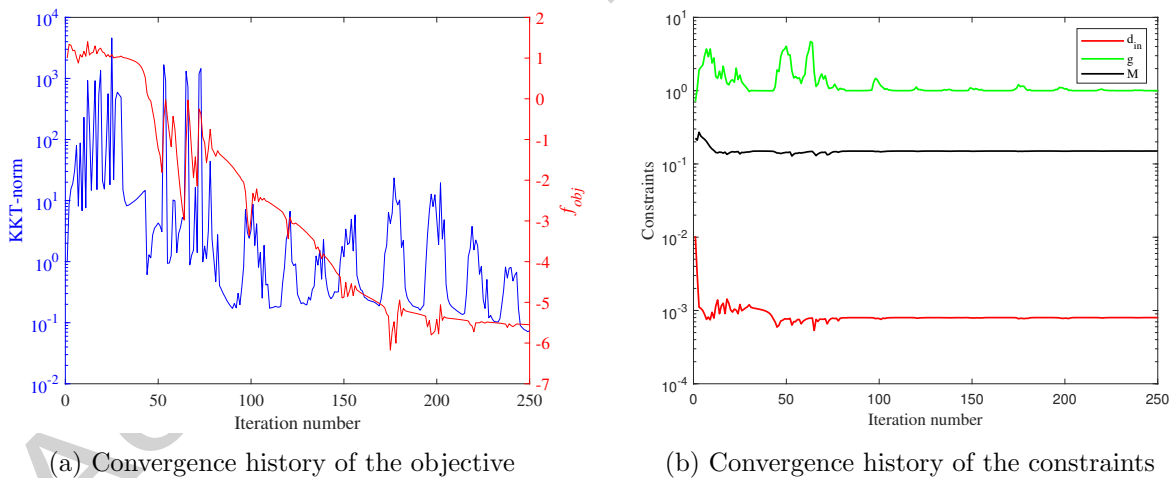
Figure 17: Convergence history of the multi-material gripper optimization

Figure 18 shows the deformation of the structure at two example time steps during the cooling process. Figure 19 shows the contribution of different thermomechanical behaviors to the displacement of point A in the vertical direction. From the results, we again observe that the transformation displacement help to increase the total tip displacement. Hence, the deformation is a result of both TWSME and superelasticity.

Figure 20 shows the variation of essential state variables of the SMA material related to the TWSME, evaluated at mesh nodes and Gauss points that are closest to sample points B, C and

26

(a) t = 0　　　　　　　　　　　　　　　　　(b) t = 20
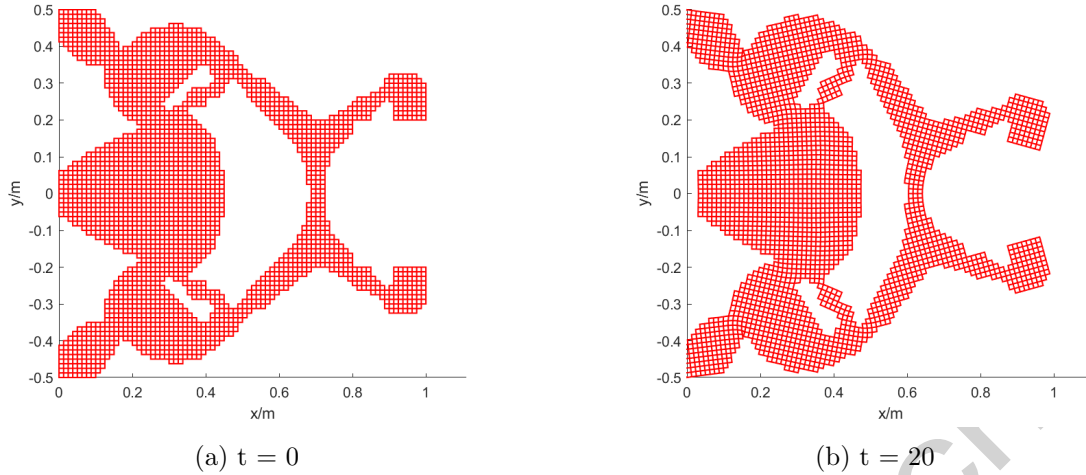
Figure 18: Example deformation of the gripper during the cooling process
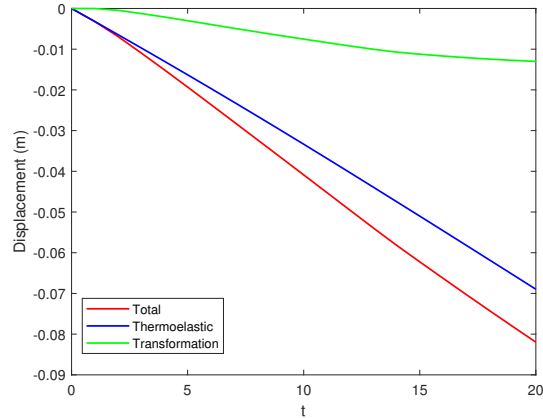
Figure 19: Contribution of different thermomechanical behaviors to the tip displacement of point A in y direction

D (see Figure 16). Again, evolution of the martensite volume fraction, horizontal transformation strain and stress-strain indicate the sequential transformation behaviors of the three points. In addition, the hysteresis of the stress-strain relationship also indicates the presence of superelasticity and TWSME.

3.3. The Impact of Considering Transient Heat Conduction

In this section, we discuss the impact of considering transient heat conduction in the proposed design framework. For the purpose of comparison, we investigate a multimaterial gripper design considering transient thermal conduction, and we compare it with a similar design produced while assuming a spatially uniform, time-varying temperature distribution. We still use the three material candidates (plus void), whose properties are stated in Tables 2 and 3. In addition, the same design

27

(a) Temperature evolution

(b) Martensite volume fraction evolution

(c) Transformation strain $\boldsymbol{\varepsilon}_x^t$ evolution

(d) Strain $\boldsymbol{\varepsilon}_x^t$-stress $\boldsymbol{\sigma}_x^t$ evolution
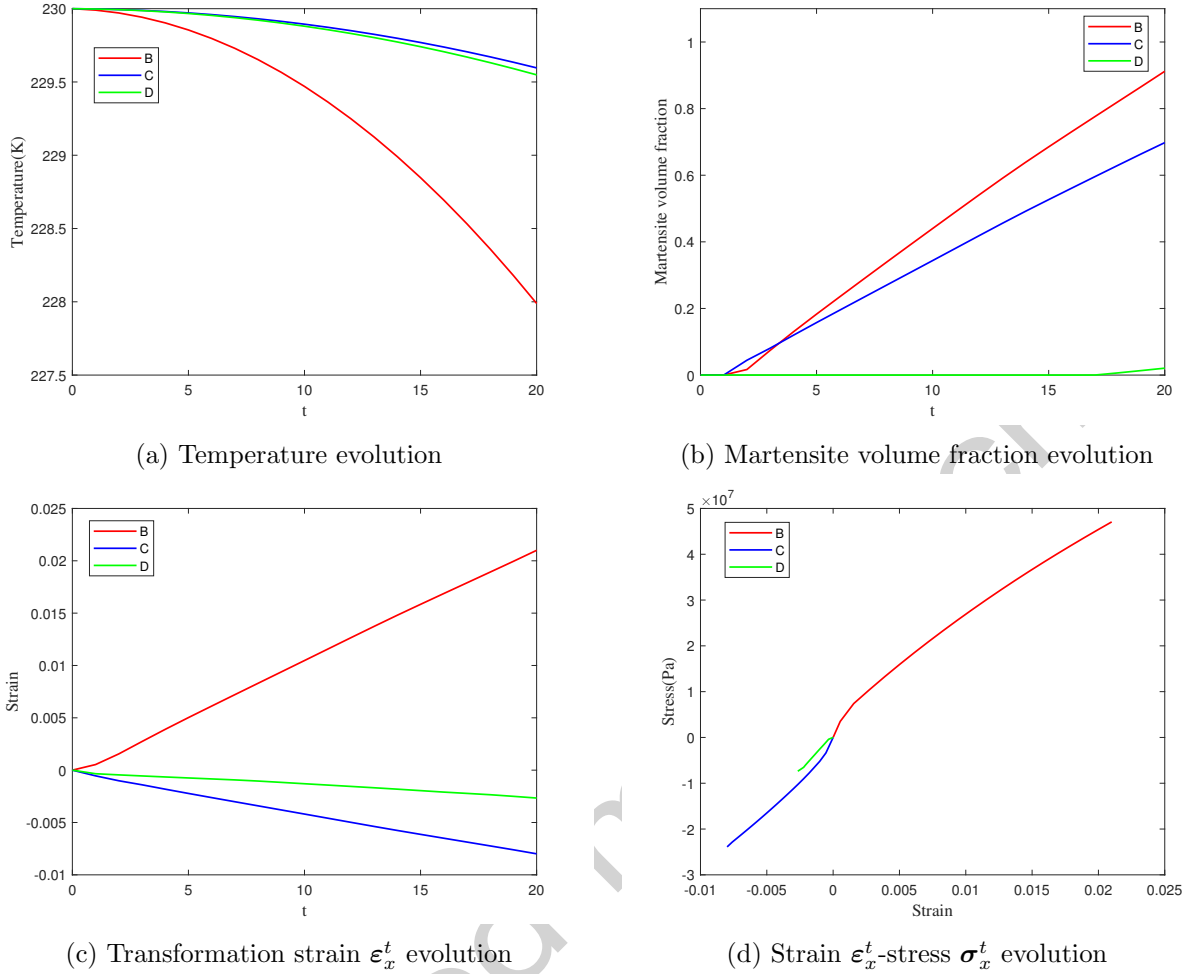
Figure 20: Evolution of the SMA state variables at three points in the structure

domain presented in Figure 15 is adopted and the optimization problem is stated in Equation 40

$$
\begin{aligned}
\min \quad & f_{obj} = -|\boldsymbol{L}^{\mathrm{T}}\boldsymbol{d}|, \\
s.t \quad & d_{in} < 2 \times 10^{-2}\text{m} \\
& M = \frac{\sum_{i=1}^{Ne}\sum_{j=2}^{4}\mu_i^{(j)}E^{(j)}}{E_{max} \cdot Ne} \le 0.12
\end{aligned}
\tag{40}
$$

where we want to maximize the vertical deformation of point A triggered by both the TWSME and superelasticity (note that we assume symmetry so we design and simulate only the top half of the mechanism). Similar to the previous problem, we have an initial displacement constraint of 2×10^{-2}m to avoid an overly compliant design. In addition, the relative mass of the structure M, cf. Equation 40, is set to be 0.12.

On the left side of the design domain, an input force increasing from 0 to $\boldsymbol{F}_{in} = 1.25 \times 10^8$N is applied in 20 time steps with a constant increment. Two springs, with stiffness $k_s = 5$N/m, are attached to the tips of the gripper. The structure is subjected to an initial temperature field of $230K$, with a reference temperature of $240K$.

Again, we use a mesh of 5888 quadrilateral elements for the finite element analysis. For the

28

design considering transient thermal conduction, the effective physical properties are calculated via Equations 22 and 24. To trigger the TWMSE, a cooling process over 20 time steps is implemented. The outward heat flux of the cooling process is set to increase incrementally to $q = 2.5 \times 10^3$ w/m, with the heat sink decreasing to a maximum of $Q = -1.25 \times 10^5$ w/m^2. For the case assuming uniform temperature distribution, we assume the temperature field is uniform across the design domain and the whole design domain is cooled with a uniform temperature increment of -0.1K. The effective physical properties representing the presence of material candidates are the Young's modulus and material density, which also follow Equation 22. For both problems, the temperature $\bar{T}$ at the left side of the domain remains at 230K during the cooling process and the penalization factor P is chosen to be 3. A density filter with a radius twice of the minimum distance between the centers of two neighboring elements is applied. The optimization results are shown in Figure 21

(a) Design considering transient heat conduction

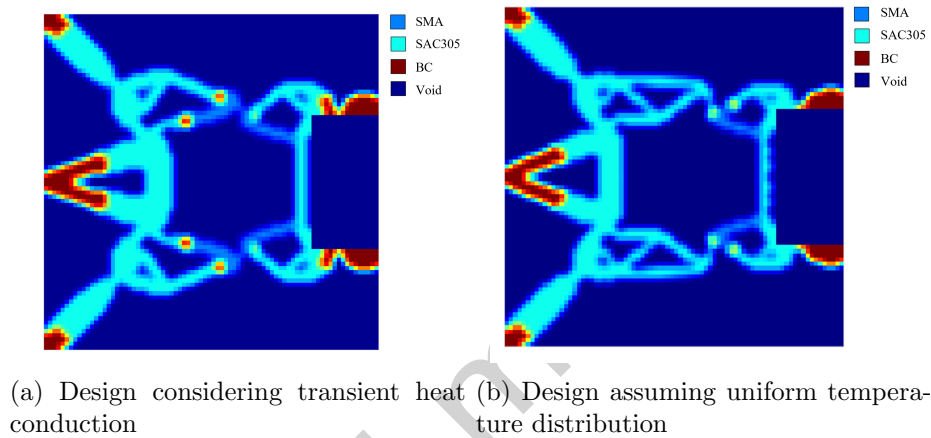(b) Design assuming uniform temperature distribution

Figure 21: Gripper designs considering different thermal conduction models

After obtaining the results, we simulate both designs for 20 time steps considering transient thermal conduction, where the boundary conditions for both cases exactly follow the those used for the topology design considering transient thermal conduction. Figure 22 compares the contribution of different thermomechanical behaviors under different design strategies to the vertical displacement of point A. From the results, we can observe that the transformation displacement has been improved by 186% by considering transient thermal conduction in the design. This suggests that considering transient thermal conduction not only improves design performance, but is also more application-oriented in considering influence of multiple material candidates on the overall thermomechanical behavior of SMA-based structures.

4. Manufacturability and Possible Applications

The designs presented above can be fabricated through the technique of multi-metal 3D printing. Manufacturing metal structures via 3D printing, especially considering multimaterial designs, is an emerging but promising research domain [50]. Possible 3D printing approaches to realize the proposed design framework could be laser powder bed fusion (LPBF) and pressure-assisted sintering, equipped with multimaterial decomposition capabilities [50, 51].

With the proposed framework, one can expand the capability and thermal performance of SMA-based smart structures to achieve better heating efficiency and bandwidth of the designed mechanisms. In addition, the proposed framework could be extended to many other application scenarios.

29
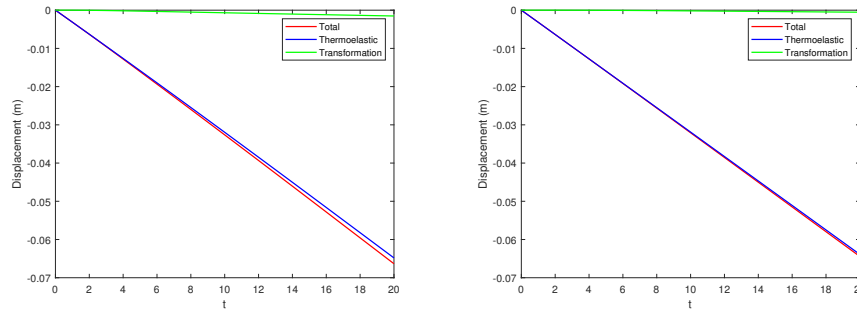
(a) Considering transient heat conduction

(b) Assuming uniform temperature distribution

Figure 22: Boundary conditions for topology optimization of a multimaterial gripper

First, constraints considering active control and manufacturing requirements could be incorporated into the proposed design framework. Such incorporation could lead to fabrication-compatible SMA-based smart structures with advanced programmability and less complexity. Second, the proposed framework could be adopted in designing a number of compliant mechanisms which rely heavily on SMAs, including smart wings, self-folding and origami mechanisms, intelligent biomedical and drug-delivery devices, and soft robotics [3].

5. Conclusions

This paper presents a novel computational design framework for SMA-based active structures with programmable morphology. The merits of the paper can be summarized in the following four aspects. 1) The proposed framework enables the selection of multiple design materials. It also takes into account the coupling effect between transient thermal conduction problem and the inelastic responses of SMAs. 2) The proposed topology optimization algorithm is able to capture changes in the transformation starting and ending criteria based on environmental conditions, which makes the design free from assumptions of fixed transformation criteria. 3) We implement a novel SIMP-like scheme (cf. Equation 24) to interpolate physical properties of the material candidates. 4) We present examples that incorporate thermal constraints in the design of SMA-based smart structures. Such constraints can be expanded to deal with various types of concerns related to the thermal performance of SMA-based structures. Case studies show that our design algorithm is applicable to both the TWSME and superelastic behaviors. Specifically, the proposed framework is able to create active bending structure triggered by the TWSME, and a gripping mechanism that utilizes both superelasticity and TWSME. Ultimately, the framework provide a viable approach to program SMA-based structures considering both their thermal and mechanical performances.

Acknowledgements

Conflict of interest

The authors declare that they have no conflict of interest.

30

Replication of results

Detailed descriptions of the algorithms used to generate all results are provided throughout the paper. Additionally, we have included all relevant material properties, and all algorithm parameters. Copies of the code used to generate the results will be made available upon request.

Appendix A. Analytical Derivation of Adjoint Vectors

In this appendix, we present the derivation of an analytical formulation for computing the tangent matrices described in Equation 35. The motivation behind deriving these formulas is to avoid directly factorizing the ill-conditioned matrix $\partial \boldsymbol{H}_n / \partial \boldsymbol{\nu}_n$. For conciseness, we focus on the case where the behaviors of the SMA in the current time step and next step are inelastic. To start, we first represent the inverse of $\partial \boldsymbol{H}_{\mathfrak{G},n} / \partial \boldsymbol{\nu}_{\mathfrak{G},n}$ using the Schur complement [52].

$$\left(\frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \right)^{-1} = \begin{bmatrix} (\mathbb{A} + \mathbb{B}\mathbb{C})^{-1} & (\mathbb{A} + \mathbb{B}\mathbb{C})^{-1}\mathbb{B} \\ \mathbb{C}(\mathbb{A} + \mathbb{B}\mathbb{C})^{-1} & -\boldsymbol{I} + \mathbb{C}(\mathbb{A} + \mathbb{B}\mathbb{C})^{-1}\mathbb{B} \end{bmatrix} \tag{A.1}$$

with

$$\frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} = \left[\begin{array}{ccc|c} \partial_\xi \Phi_n & \boldsymbol{0}_{1\times 6} & 0_{1\times 1} & \partial_{\boldsymbol{\sigma}_n} \Phi_n^{\mathrm{T}} \\ \boldsymbol{\Lambda}_n & -\boldsymbol{I}_{6\times 6} & \boldsymbol{0}_{6\times 1} & \partial_{\boldsymbol{\sigma}} \boldsymbol{\Lambda}_n : \Delta \xi_n \\ \Delta S & \boldsymbol{0}_{1\times 6} & -I_{1\times 1} & \boldsymbol{0}_{1\times 6} \\ \hline \boldsymbol{0}_{6\times 1} & -\boxed{\boldsymbol{S}_n^{-1}} & -\boxed{S_n^{-1}} : \boldsymbol{\sigma}_n & -\boldsymbol{I}_{6\times 6} \end{array} \right] = \begin{bmatrix} \mathbb{A} & \mathbb{B} \\ \mathbb{C} & \mathbb{D} \end{bmatrix} \tag{A.2}$$

Note that the analytical solution in which we directly factorize $\partial \boldsymbol{H}_n / \partial \boldsymbol{\nu}_n$ is not recommended for calculating the sensitivities. This analytical solution will not change the ill-conditioned characteristics of the matrix, since $(\mathbb{A} + \mathbb{B}\mathbb{C})$ is nearly singular due to the fact that the large value S^{-1} still appears on the lower triangle. Similarly, $\partial \boldsymbol{H}_{n+1} / \partial \boldsymbol{\nu}_n$ can be defined in block matrix form as

$$\frac{\partial \boldsymbol{H}_{n+1}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} = \left[\begin{array}{ccc|c} 0_{1\times 1} & \boldsymbol{0}_{1\times 6} & 0_{1\times 1} & \boldsymbol{0}_{1\times 6} \\ -\boldsymbol{\Lambda}_{n+1} & \boldsymbol{I}_{6\times 6} & \boldsymbol{0}_{6\times 1} & \boldsymbol{0}_{6\times 6} \\ -\Delta S & \boldsymbol{0}_{1\times 6} & I_{1\times 1} & \boldsymbol{0}_{1\times 6} \\ \hline \boldsymbol{0}_{6\times 1} & \boldsymbol{0}_{6\times 6} & \boldsymbol{0}_{6\times 1} & \boldsymbol{0}_{6\times 6} \end{array} \right] = \begin{bmatrix} \widetilde{\mathbb{A}} & \widetilde{\mathbb{B}} \\ \widetilde{\mathbb{C}} & \widetilde{\mathbb{D}} \end{bmatrix} \tag{A.3}$$

The five aforementioned matrices then can be represented in a new form shown below.

$$-\bigwedge_{\mathrm{el}} \sum_{\mathfrak{G}} \frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \left(\frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \right)^{-1} \frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{u}_{\mathrm{el},n}} = \bigwedge_{\mathrm{el}} \sum_{\mathfrak{G}} w \boldsymbol{B}_{\mathfrak{G}}^{\mathrm{T}} [\boldsymbol{I} - \mathbb{C}(\mathbb{A} + \mathbb{B}\mathbb{C})^{-1}\mathbb{B}]_{\mathfrak{G}} \boldsymbol{S}_{\mathfrak{G},n}^{-1} \boldsymbol{B}_{\mathfrak{G}} \det \boldsymbol{J}_{\mathfrak{G}}$$

$$\frac{\partial \boldsymbol{R}_{\mathrm{el},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \right)^{-1} = w \boldsymbol{B}^{\mathrm{T}} \left[\begin{array}{cc} \mathbb{C}(\mathbb{A} + \mathbb{B}\mathbb{C})^{-1} & -\boldsymbol{I} + \mathbb{C}(\mathbb{A} + \mathbb{B}\mathbb{C})^{-1}\mathbb{B} \end{array} \right] \det \boldsymbol{J}$$

$$\left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \right)^{-1} \frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{u}_{\mathrm{el},n}} = \left[\begin{array}{c} (\mathbb{A} + \mathbb{B}\mathbb{C})^{-1}\mathbb{B} \\ -I + \mathbb{C}(\mathbb{A} + \mathbb{B}\mathbb{C})^{-1}\mathbb{B} \end{array} \right] \boldsymbol{B} \tag{A.4}$$

$$\frac{\partial \boldsymbol{H}_{\mathfrak{G},n+1}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \right)^{-1} = \left[\begin{array}{cc} \widetilde{\mathbb{A}}(\mathbb{A} + \mathbb{B}\mathbb{C})^{-1} & \widetilde{\mathbb{A}}(\mathbb{A} + \mathbb{B}\mathbb{C})^{-1}\mathbb{B} \\ (\widetilde{\mathbb{C}} + \mathbb{C})(\mathbb{A} + \mathbb{B}\mathbb{C})^{-1} & -\boldsymbol{I} + (\widetilde{\mathbb{C}} + \mathbb{C})(\mathbb{A} + \mathbb{B}\mathbb{C})^{-1}\mathbb{B} \end{array} \right]$$

$$\frac{\partial \boldsymbol{H}_{\mathfrak{G},n+1}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},\mathfrak{G},n}} \right)^{-1} \frac{\partial \boldsymbol{H}_n}{\partial \boldsymbol{u}_{\mathrm{el},n}} = \left[\begin{array}{c} \widetilde{\mathbb{A}}(\mathbb{A} + \mathbb{B}\mathbb{C})^{-1}\mathbb{B} - \boldsymbol{I} \end{array} \right] \boldsymbol{S}_n^{-1} \boldsymbol{B}$$

31

It can be observed that the key to obtaining an accurate solution of the tangent matrices lies in solving the inverse of $(\mathbb{A} + \mathbb{B}\mathbb{C})$. Here we use the Schur formulation again to calculate the inverse of the matrix.

$$(\mathbb{A} + \mathbb{B}\mathbb{C})^{-1} = \begin{bmatrix} (\mathbb{A}' - \mathbb{B}'\mathbb{D}'^{-1}\mathbb{C}')^{-1} & -(\mathbb{A}' - \mathbb{B}'\mathbb{D}'^{-1}\mathbb{C}')^{-1}\mathbb{B}'\mathbb{D}'^{-1} \\ -\mathbb{D}'^{-1}\mathbb{C}'(\mathbb{A}' - \mathbb{B}'\mathbb{D}'^{-1}\mathbb{C}')^{-1} & \mathbb{D}'^{-1} + \mathbb{D}'^{-1}\mathbb{C}'(\mathbb{A}' - \mathbb{B}'\mathbb{D}'^{-1}\mathbb{C}')^{-1}\mathbb{B}'\mathbb{D}'^{-1} \end{bmatrix} \quad (A.5)$$

with

$$\mathbb{A} + \mathbb{B}\mathbb{C} = \left[\begin{array}{c|cc} \partial_\xi \Phi_n & -\partial_{\boldsymbol{\sigma}}\Phi_n : \boldsymbol{S}_n^{-1} & -\partial_{\boldsymbol{\sigma}}\Phi_n : S_n^{-1} : \boldsymbol{\sigma}_n \\ \hline \boldsymbol{\Lambda}_n & -(\boldsymbol{S}_n + \partial_{\boldsymbol{\sigma}}\boldsymbol{\Lambda}_n : \Delta\xi_n) : \boldsymbol{S}_n^{-1} & -\partial_{\boldsymbol{\sigma}}\boldsymbol{\Lambda}_n : \Delta\xi_n : S_n^{-1} : \boldsymbol{\sigma}_n \\ \Delta S & \boldsymbol{0}_{1\times 6} & -I_{1\times 1} \end{array} \right] = \begin{bmatrix} \mathbb{A}' & \mathbb{B}' \\ \mathbb{C}' & \mathbb{D}' \end{bmatrix} \quad (A.6)$$

Here, $\mathbb{D}'^{-1}$ and $(\mathbb{A}' - \mathbb{B}'\mathbb{D}'^{-1}\mathbb{C}')^{-1}$ need to be solved. Note that it can be easily proven that $\partial_{\boldsymbol{\sigma}}\boldsymbol{\Lambda}$ belongs to the null space of $\boldsymbol{\sigma}$, i.e. $\partial_{\boldsymbol{\sigma}}\boldsymbol{\Lambda}:\boldsymbol{\sigma} = 0$, hence $\partial_{\boldsymbol{\sigma}}\boldsymbol{\Lambda}$ is not invertible. However, $(\boldsymbol{S}_n + \partial_{\boldsymbol{\sigma}}\boldsymbol{\Lambda}_n : \Delta\xi_n)$ is nonsingular. Defining $\boldsymbol{\zeta} = \boldsymbol{S}_n + \partial_{\boldsymbol{\sigma}}\boldsymbol{\Lambda}_n : \Delta\xi_n$, $\mathbb{D}'^{-1}$ can be solved using the Schur form as

$$\mathbb{D}'^{-1} = \begin{bmatrix} -\boldsymbol{S}_n : \boldsymbol{\zeta}_n^{-1} & \boldsymbol{S}_n : \boldsymbol{\zeta}_n^{-1} : \partial_{\boldsymbol{\sigma}}\boldsymbol{\Lambda}_n : \Delta\xi_n : S_n^{-1} : \boldsymbol{\sigma}_n \\ \boldsymbol{0}_{1\times 6} & -I_{1\times 1} \end{bmatrix} \quad (A.7)$$

With the above information, one can obtain that

$$\begin{aligned} \mathbb{B}'\mathbb{D}'^{-1} &= \begin{bmatrix} \partial_{\boldsymbol{\sigma}}\Phi_n : \boldsymbol{\zeta}_n^{-1} & \partial_{\boldsymbol{\sigma}}\Phi_n : \boldsymbol{\zeta}_n^{-1} : \boldsymbol{S}_n^{-1} : S_n^{-1} : \boldsymbol{\sigma}_n \end{bmatrix} \\ \mathbb{D}'^{-1}\mathbb{C}' &= \begin{bmatrix} -\boldsymbol{S}_n : \boldsymbol{\zeta}_n^{-1} : (\boldsymbol{\Lambda}_n - \partial_{\boldsymbol{\sigma}}\boldsymbol{\Lambda}_n : \Delta\xi_n : S_n^{-1} : \boldsymbol{\sigma}_n : \Delta S) \\ -\Delta S \end{bmatrix} = \begin{bmatrix} -\widetilde{\boldsymbol{\Lambda}}_n \\ -\Delta S \end{bmatrix} \end{aligned} \quad (A.8)$$

Then $(\mathbb{A}' - \mathbb{B}'\mathbb{D}'^{-1}\mathbb{C}')^{-1}$ is a scalar and can be calculated as $Q = -\partial_{\boldsymbol{\sigma}}\Phi_n : \boldsymbol{\zeta}_n^{-1} : \partial_{\boldsymbol{\sigma}}\Phi_n + \partial_\xi\Phi_n$. Hence the issue of the ill-conditioned matrix has been solved, and we are able to accurately evaluate the five matrices as follows.

$$\begin{aligned} -\bigwedge_{\text{el}}\sum_{\mathfrak{G}} \frac{\partial \boldsymbol{R}_n}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \right)^{-1} \frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{u}_{\text{el},n}} &= \bigwedge_{\text{el}}\sum_{\mathfrak{G}} w\boldsymbol{B}_{\mathfrak{G}}^{\mathrm{T}}(\boldsymbol{\zeta}_{\mathfrak{G},n}^{-1} - \frac{\boldsymbol{\zeta}_{\mathfrak{G},n}^{-1} : \partial_{\boldsymbol{\sigma}}\Phi_{\mathfrak{G},n} \otimes \boldsymbol{\zeta}_{\mathfrak{G},n}^{-1} : \partial_{\boldsymbol{\sigma}}\Phi_{\mathfrak{G},n}}{\partial_{\boldsymbol{\sigma}}\Phi_{\mathfrak{G},n} : \boldsymbol{\zeta}_{\mathfrak{G},n}^{-1} : \partial_{\boldsymbol{\sigma}}\Phi_{\mathfrak{G},n} - \partial_\xi\Phi_{\mathfrak{G},n}})\boldsymbol{B}_{\mathfrak{G}}\det\boldsymbol{J}_{\mathfrak{G}} \\ &= \bigwedge_{\text{el}}\sum_{\mathfrak{G}} w\boldsymbol{B}_{\mathrm{G}}^{\mathrm{T}}\boldsymbol{\mathfrak{L}}_{\mathfrak{G},n}\boldsymbol{B}_{\mathfrak{G}}\det\boldsymbol{J}_{\mathfrak{G}} \\ \frac{\partial \boldsymbol{R}_{\text{el},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \right)^{-1} &= w\boldsymbol{B}^{\mathrm{T}} \begin{bmatrix} -\frac{\boldsymbol{\zeta}_n^{-1} : \partial_{\boldsymbol{\sigma}}\Phi_n}{Q} & \boldsymbol{\mathfrak{L}}_n & \boldsymbol{\mathfrak{L}}_n : \boldsymbol{\mathfrak{C}}^{-1} : \boldsymbol{\sigma}_n & -\boldsymbol{\mathfrak{L}}_n : \boldsymbol{S}_n \end{bmatrix} \det\boldsymbol{J} \\ \left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \right)^{-1} \frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{u}_{\text{el},n}} &= \begin{bmatrix} \frac{\partial_{\boldsymbol{\sigma}}\Phi_n^{\mathrm{T}} : \boldsymbol{\zeta}_n^{-1}}{Q} \\ \frac{\widetilde{\boldsymbol{\Lambda}}_n : \partial_{\boldsymbol{\sigma}}\Phi_n^{\mathrm{T}}}{Q} : \boldsymbol{\zeta}_n^{-1} + \boldsymbol{S}_n : \boldsymbol{\zeta}_n^{-1} - \boldsymbol{I}_{6\times 6} \\ \frac{\Delta S : \partial_{\boldsymbol{\sigma}}\Phi_n^{\mathrm{T}}}{Q} : \boldsymbol{\zeta}_n^{-1} \\ -\boldsymbol{\mathfrak{L}}_n \end{bmatrix} \boldsymbol{B} \\ \frac{\partial \boldsymbol{H}_{\mathfrak{G},n+1}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \right)^{-1} &= \begin{bmatrix} \mathfrak{A}_{8\times 8} & \mathfrak{B}_{8\times 6} \\ \boldsymbol{0}_{6\times 8} & \boldsymbol{0}_{6\times 6} \end{bmatrix} \\ \frac{\partial \boldsymbol{H}_{\mathfrak{G},n+1}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \left(\frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{\nu}_{\mathfrak{G},n}} \right)^{-1} \frac{\partial \boldsymbol{H}_{\mathfrak{G},n}}{\partial \boldsymbol{u}_{\text{el},n}} &= \begin{bmatrix} \boldsymbol{0}_{1\times 6} \\ \frac{\widetilde{\boldsymbol{\Lambda}}_n - \boldsymbol{\Lambda}_{n+1}}{Q} : \partial_{\boldsymbol{\sigma}}\Phi_n^{\mathrm{T}} : \boldsymbol{\zeta}_n^{-1} + \boldsymbol{S}_n : \boldsymbol{\zeta}_n^{-1} - \boldsymbol{I}_{6\times 6} \\ \boldsymbol{0}_{7\times 6} \end{bmatrix} \boldsymbol{B} \end{aligned} \quad (A.9)$$

32

where

$$\mathfrak{A} = \begin{bmatrix} 0_{1\times1} & \mathbf{0}_{1\times7} \\ \frac{\widetilde{\mathbf{\Lambda}}_n - \mathbf{\Lambda}_{n+1}}{Q} & \mathfrak{a}_{7\times7} \\ 0_{1\times1} & \end{bmatrix}$$

$$\mathfrak{a} = \begin{bmatrix} (-\boldsymbol{S}_n - \frac{\widetilde{\mathbf{\Lambda}}_n - \mathbf{\Lambda}_{n+1}}{Q} : \partial_{\boldsymbol{\sigma}}\Phi_n{}^{\mathrm{T}}) : \boldsymbol{\zeta}_n^{-1} & (-\frac{\widetilde{\mathbf{\Lambda}}_n - \mathbf{\Lambda}_{n+1}}{Q} : \partial_{\boldsymbol{\sigma}}\Phi_n{}^{\mathrm{T}} : \boldsymbol{\zeta}_n^{-1} : \boldsymbol{S}_n + \boldsymbol{S}_n : \boldsymbol{\zeta}_n^{-1} : \partial_{\boldsymbol{\sigma}}\mathbf{\Lambda}_n : \Delta\xi_n) : S_n^{-1} : \boldsymbol{\sigma}_n \\ \mathbf{0}_{1\times6} & -I_{1\times1} \end{bmatrix}$$

$$\mathfrak{B} = \begin{bmatrix} \mathbf{0}_{1\times6} \\ \frac{\widetilde{\mathbf{\Lambda}}_n - \mathbf{\Lambda}_{n+1}}{Q} : \partial_{\boldsymbol{\sigma}}\Phi_n{}^{\mathrm{T}} : \boldsymbol{\zeta}_n^{-1} + \boldsymbol{S}_n : \boldsymbol{\zeta}_n^{-1} - \boldsymbol{I}_{6\times6} \\ \mathbf{0}_{1\times6} \end{bmatrix} : \boldsymbol{S}_n \qquad (\text{A.10})$$

$$\widetilde{\mathbf{\Lambda}}_n = \boldsymbol{S}_n : \boldsymbol{\zeta}_n^{-1} : (\mathbf{\Lambda}_n - \partial_{\boldsymbol{\sigma}}\mathbf{\Lambda}_n : \Delta\xi_n : S_n^{-1} : \boldsymbol{\sigma}_n : \Delta S)$$

When $\partial_{\boldsymbol{\sigma}}\mathbf{\Lambda} = 0$ and $\boldsymbol{\zeta} = \boldsymbol{S}$, the above analytical solutions degenerate as follows

$$-\bigwedge_{\mathrm{el}}\sum_{\mathfrak{G}}\frac{\partial\boldsymbol{R}_n}{\partial\boldsymbol{\nu}_{\mathfrak{G},n}}\left(\frac{\partial\boldsymbol{H}_{\mathfrak{G},n}}{\partial\boldsymbol{\nu}_{\mathfrak{G},n}}\right)^{-1}\frac{\partial\boldsymbol{H}_{\mathfrak{G},n}}{\partial\boldsymbol{u}_{\mathrm{el},n}} = \bigwedge_{\mathrm{el}}\sum_{\mathfrak{G}}w\boldsymbol{B}_{\mathfrak{G}}^{\mathrm{T}}(\boldsymbol{S}_{\mathfrak{G},n}^{-1} - \frac{\boldsymbol{S}_{\mathfrak{G},n}^{-1} : \partial_{\boldsymbol{\sigma}}\Phi_{\mathfrak{G},n} \otimes \boldsymbol{S}_{\mathfrak{G},n}^{-1} : \partial_{\boldsymbol{\sigma}}\Phi_{\mathfrak{G},n}}{\partial_{\boldsymbol{\sigma}}\Phi_{\mathfrak{G},n} : \boldsymbol{S}_{\mathfrak{G},n}^{-1} : \partial_{\boldsymbol{\sigma}}\Phi_{\mathfrak{G},n} - \partial_\xi\Phi_{\mathfrak{G},n}})\boldsymbol{B}_{\mathfrak{G}}\det\boldsymbol{J}_{\mathfrak{G}}$$

$$= \bigwedge_{\mathrm{el}}\sum_{\mathfrak{G}}w\boldsymbol{B}_{\mathfrak{G}}^{\mathrm{T}}\mathfrak{L}_{\mathfrak{G},n}\boldsymbol{B}_{\mathfrak{G}}\det\boldsymbol{J}_{\mathfrak{G}}$$

$$\frac{\partial\boldsymbol{R}_{\mathrm{el},n}}{\partial\boldsymbol{\nu}_{\mathfrak{G},n}}\left(\frac{\partial\boldsymbol{H}_{\mathfrak{G},n}}{\partial\boldsymbol{\nu}_{\mathfrak{G},n}}\right)^{-1} = w\boldsymbol{B}^{\mathrm{T}}\begin{bmatrix} \frac{\boldsymbol{S}_n^{-1}:\partial_{\boldsymbol{\sigma}}\Phi_n}{Q} & \mathfrak{L}_n & \mathfrak{L}_n : \mathfrak{C}^{-1} : \boldsymbol{\sigma}_n & -\mathfrak{L}_n : \boldsymbol{S}_n \end{bmatrix}\det\boldsymbol{J}$$

$$\left(\frac{\partial\boldsymbol{H}_{\mathfrak{G},n}}{\partial\boldsymbol{\nu}_{\mathfrak{G},n}}\right)^{-1}\frac{\partial\boldsymbol{H}_{\mathfrak{G},n}}{\partial\boldsymbol{u}_{\mathrm{el},n}} = \begin{bmatrix} \frac{\partial_{\boldsymbol{\sigma}}\Phi_n{}^{\mathrm{T}}:\boldsymbol{S}_n^{-1}}{Q} \\ \frac{\begin{bmatrix}\mathbf{\Lambda}_n \\ \Delta S\end{bmatrix}:\partial_{\boldsymbol{\sigma}}\Phi_n{}^{\mathrm{T}}:\boldsymbol{S}_n^{-1}}{Q} \\ -\mathfrak{L}_n \end{bmatrix}\boldsymbol{B} \qquad (\text{A.11})$$

$$\frac{\partial\boldsymbol{H}_{\mathfrak{G},n+1}}{\partial\boldsymbol{\nu}_{\mathfrak{G},n}}\left(\frac{\partial\boldsymbol{H}_{\mathfrak{G},n}}{\partial\boldsymbol{\nu}_{\mathfrak{G},n}}\right)^{-1} = \begin{bmatrix} \mathfrak{A}_{8\times8} & \mathfrak{B}_{8\times6} \\ \mathbf{0}_{6\times8} & \mathbf{0}_{6\times6} \end{bmatrix}$$

$$\frac{\partial\boldsymbol{H}_{\mathfrak{G},n+1}}{\partial\boldsymbol{\nu}_{\mathfrak{G},n}}\left(\frac{\partial\boldsymbol{H}_{\mathfrak{G},n}}{\partial\boldsymbol{\nu}_{\mathfrak{G},n}}\right)^{-1}\frac{\partial\boldsymbol{H}_{\mathfrak{G},n}}{\partial\boldsymbol{u}_{\mathrm{el},n}} = \begin{bmatrix} \mathbf{0}_{1\times6} \\ \left(\frac{\mathbf{\Lambda}_n - \mathbf{\Lambda}_{n+1}}{Q} : \partial_{\boldsymbol{\sigma}}\Phi_n{}^{\mathrm{T}} : \boldsymbol{S}_n^{-1}\right) \\ \mathbf{0}_{7\times6} \end{bmatrix}\boldsymbol{B}$$

where

$$\mathfrak{A} = \begin{bmatrix} 0_{1\times1} & \mathbf{0}_{1\times7} \\ \frac{\mathbf{\Lambda}_n - \mathbf{\Lambda}_{n+1}}{Q} & -\boldsymbol{I}_{7\times7} - \frac{\mathbf{\Lambda}_n - \mathbf{\Lambda}_{n+1}}{Q} : \partial_{\boldsymbol{\sigma}}\Phi_n{}^{\mathrm{T}} : \begin{bmatrix}\boldsymbol{S}_n^{-1} & \boldsymbol{S}_n^{-1} : \boldsymbol{\sigma}_n \\ \mathbf{0}_{1\times6} & 0_{1\times1}\end{bmatrix} \\ 0_{1\times1} & \end{bmatrix}$$

$$\mathfrak{B} = \begin{bmatrix} \mathbf{0}_{1\times6} \\ \frac{\mathbf{\Lambda}_n - \mathbf{\Lambda}_{n+1}}{Q} : \partial_{\boldsymbol{\sigma}}\Phi_n{}^{\mathrm{T}} \\ \mathbf{0}_{1\times6} \end{bmatrix} \qquad (\text{A.12})$$

Note that the first equation in A.9 and A.11 can be recognized as the formulas for calculating the tangent stiffness matrix, shown in our work [43]. Therefore, in addition to improving the accuracy of the sensitivity evaluation, we are able to conserve computational resources by re-using the tangent stiffness matrix already computed during the finite element analysis.

33

Appendix B. Comparison between Analytical and Finite Difference Sensitivity Results

This section validates the adjoint sensitivity formulation by comparing the analytical results with finite differences. Figures B.23 and B.24 show the geometry and boundary conditions of the test case used to calculate the sensitivities in Figure 6. Note that in the following tables from Table B.4 to Table B.9, t refers to the time step; T and ξ refer to the temperature and martensite volume fraction evaluated at the node and Gauss point that are closest to point A, which is associated with the function of interest f_{int}. ANA and FD refer to the analytical and finite difference sensitivity values, respectively.

Appendix B.1. Simulating Two-Way shape Memory Effects

In the first case study we evaluate the sensitivity of the TWSME of a bi-material beam containing SMA and BC, whose properties are shown in Tables 2 and 3. The material volume fraction r (cf. Equation 23) is uniformly 0.9 across the structure. The boundary conditions are shown below,
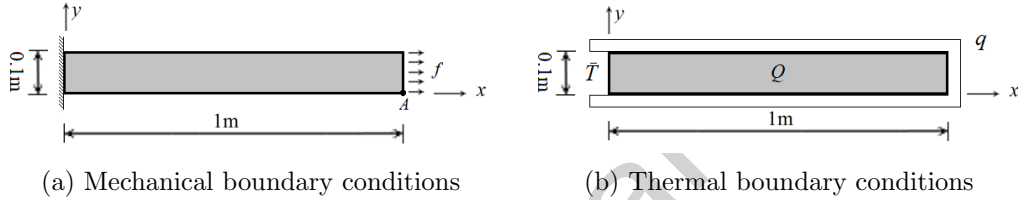

(a) Mechanical boundary conditions   (b) Thermal boundary conditions

Figure B.23: Boundary conditions for topology optimization of an active bending beam

where a constant uniform surface traction $f = 9 \times 10^7 \text{N/m}$ is applied when cooling the beam. The initial and reference temperature of the beam is set to be 250K. In addition, temperature $\bar{T}$ (cf. Figure B.23b) is fixed at 250K during the cooling process. A heat flux increment $dq = 1 \times 10^2$ w/m per time step, and a heat sink increment $dQ = -5 \times 10^3$ w/m^2 per time step are applied to cool the beam. The function of interest f_{int} for sensitivity analysis (cf. Equation 31) is the vertical displacement at point A of the structure, i.e. $f_{int} = d_y^A$. We calculate the sensitivity of the displacement function under different mesh sizes and at different transformation stages. The numerical results are shown in Tables B.4 to B.6.

Value	Mesh size = 4				
t	10	50	100	150	170
T	249.6703199013790	243.4108661937840	226.6989523019790	201.2498689099600	188.6956589995470
ξ	0.0384723828813	0.1257331959114	0.3982380590771	0.8550579710711	1.0000000000000
ANA	0.0000742568780	0.0001051655000	0.0001464651100	0.0000242208480	-0.0009959078800
FD	0.0000742567390	0.0001051634100	0.0001464646900	0.0000242194900	-0.0009959392900

Table B.4: Sensitivity analysis when simulating TWSME, mesh size = 4

Value	Mesh size = 40				
t	10	50	100	120	150
T	249.6788491606810	243.9646147270020	229.9261144923640	222.6420380278410	210.1483510818130
ξ	0.0309148427582	0.1490178737523	0.4948131346680	0.7080530497616	1.0000000000000
ANA	-0.0021818071000	-0.0042628965000	-0.0110670770000	-0.0150625560000	-0.0223159050000
FD	-0.0021817387000	-0.0042625358000	-0.0110673980000	-0.0150628300000	-0.0223164680000

Table B.5: Sensitivity analysis when simulating TWSME, mesh size = 40

34

Value	Mesh size = 250				
t	10	50	100	120	150
T	249.7196231887840	244.2936841576180	230.3772679138910	222.8385017443770	208.9956210947780
ξ	0.0375538860204	0.1510085675313	0.5003307443236	0.6994255087386	0.9882110911420
ANA	-0.0024394635000	-0.0052037250000	-0.0112467850000	-0.0144570200000	-0.0199850280000
FD	-0.0024393834000	-0.0052038126000	-0.0112464380000	-0.0144576100000	-0.0199847610000

Table B.6: Sensitivity analysis when simulating TWSME, mesh size = 250

Appendix B.2. Simulating both Two-Way Shape Memory Effects and Superelasticity

The second case study looks at sensitivity analysis of a single-material SMA structure undergoing both TWSME and superelasticity. The material volume fraction r (cf. Equation 23) is uniformly 0.5 across the structure. The boundary conditions are shown in Figure B.24. An input force increment $d\boldsymbol{F}_{in} = 1 \times 10^6$N per time step is applied to the structure to trigger superelasticity. Meanwhile, the structure is cooled with a heat flux increment $dq = 1 \times 10^2$ w/m per time step, and a heat sink increment $dQ = -5 \times 10^3$ w/m^2 per time step. The initial and reference temperature, as well as the fixed temperature $\bar{T}$ for the thermal conduction problem (cf. Figure B.24b) are set to be $230K$. The function of interest f_{int} for sensitivity analysis (cf. Equation 31) is the mechanical advantage at point A, i.e. $f_{int} = -\boldsymbol{F}_{out}/\boldsymbol{F}_{in}$. We calculate the sensitivity of the structure, under different mesh sizes and at different transformation stages. The numerical results are shown in Tables B.7 to B.9
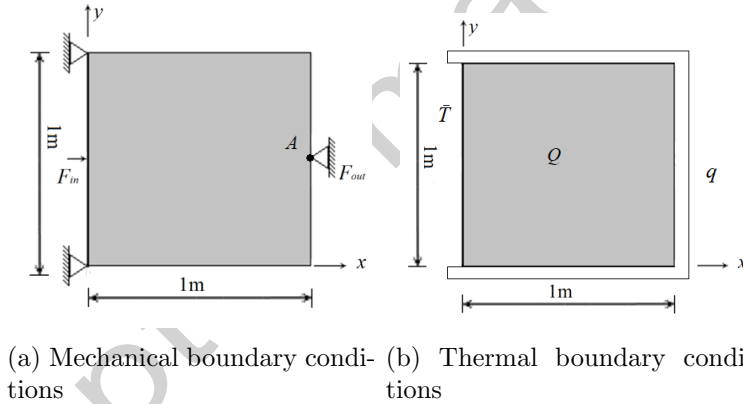

(a) Mechanical boundary conditions

(b) Thermal boundary conditions

Figure B.24: Boundary conditions for topology optimization of a force inverter

Value	Mesh size = 8				
t	10	20	30	40	50
T	229.9817293845000	229.1483928459300	227.9873025880070	225.5063204155040	222.0468478279230
ξ	0.0049697403094	0.4315482261951	0.7293862033743	0.9729830835475	1.0000000000000
ANA	0.0069215218000	0.0188556060000	0.0653826710000	0.0623086300000	0.1883299100000
FD	0.0069213230000	0.0188556530000	0.0653823980000	0.0623083090000	0.1883240900000

Table B.7: Sensitivity analysis when simulating both TWSME and superelasticity, mesh size = 8

[1] H. Rodrigue, W. Wang, M.-W. Han, T. J. Kim, and S.-H. Ahn, "An overview of shape memory alloy-coupled actuators and robots," *Soft robotics*, vol. 4, no. 1, pp. 3–15, 2017.

[2] S. Kim, C. Laschi, and B. Trimmer, "Soft robotics: a bioinspired evolution in robotics," *Trends in biotechnology*, vol. 31, no. 5, pp. 287–294, 2013.

35

Value	Mesh size = 100				
t	2	4	6	8	10
T	229.9993533215760	229.9980838273750	229.9962135196510	229.9937639881610	229.9907564165590
ξ	0.0588513902329	0.2682254842918	0.4803767358099	0.6928805383742	0.9084259171683
ANA	0.0264168500000	0.0689047670000	0.0677846290000	0.0684547530000	0.0680963820000
FD	0.0264168300000	0.0689046710000	0.0677832850000	0.0684545220000	0.0680994980000

Table B.8: Sensitivity analysis when simulating both TWSME and superelasticity, mesh size = 100

Value	Mesh size = 400				
t	2	4	6	8	10
T	229.9922613718580	229.9768809768750	229.9539472465000	229.5132938476400	229.9235462016260
ξ	0.1232803996033	0.3794228002526	0.5931090525297	0.8454409188559	1.0000000000000
ANA	0.0454707150000	0.0482810380000	0.0548983200000	0.0624812240000	0.0718520280000
FD	0.0454730150000	0.0482870690000	0.0548988279000	0.0624810190000	0.0718529390000

Table B.9: Sensitivity analysis when simulating both TWSME and superelasticity, mesh size = 400

[3] J. M. Jani, M. Leary, A. Subic, and M. A. Gibson, "A review of shape memory alloy research, applications and opportunities," *Materials & Design (1980-2015)*, vol. 56, pp. 1078–1113, 2014.

[4] M. H. Elahinia, *Shape memory alloy actuators: design, fabrication, and experimental evaluation.* John Wiley & Sons, 2016.

[5] D. C. Lagoudas, *Shape memory alloys: modeling and engineering applications.* Springer, 2008.

[6] Q. P. Sun and K. C. Hwang, "Micromechanics modelling for the constitutive behavior of polycrystalline shape memory alloys—i. derivation of general relations," *Journal of the Mechanics and Physics of Solids*, vol. 41, no. 1, pp. 1–17, 1993.

[7] Q. P. Sun and K. C. Hwang, "Micromechanics modelling for the constitutive behavior of polycrystalline shape memory alloys—ii. study of the individual phenomena," *Journal of the Mechanics and Physics of Solids*, vol. 41, no. 1, pp. 19–33, 1993.

[8] K. Bhattacharya and R. V. Kohn, "Symmetry, texture and the recoverable strain of shape-memory polycrystals," *Acta materialia*, vol. 44, no. 2, pp. 529–542, 1996.

[9] K. Bhattacharya *et al.*, *Microstructure of martensite: why it forms and how it gives rise to the shape-memory effect*, vol. 2. Oxford University Press, 2003.

[10] C. Cisse, W. Zaki, and T. B. Zineb, "A review of constitutive models and modeling techniques for shape memory alloys," *International Journal of Plasticity*, vol. 76, pp. 244–284, 2016.

[11] M. Frost, B. Benešová, and P. Sedlák, "A microscopically motivated constitutive model for shape memory alloys: formulation, analysis and computations," *Mathematics and Mechanics of Solids*, vol. 21, no. 3, pp. 358–382, 2016.

[12] K. Tanaka, F. Nishimura, T. Hayashi, H. Tobushi, and C. Lexcellent, "Phenomenological analysis on subloops and cyclic behavior in shape memory alloys under mechanical and/or thermal loads," *Mechanics of Materials*, vol. 19, no. 4, pp. 281–292, 1995.

[13] C. Liang and C. Rogers, "A multi-dimensional constitutive model for shape memory alloys," *Journal of Engineering Mathematics*, vol. 26, no. 3, pp. 429–443, 1992.

36

[14] C. Liang and C. A. Rogers, "One-dimensional thermomechanical constitutive relations for shape memory materials," *Journal of intelligent material systems and structures*, vol. 8, no. 4, pp. 285–302, 1997.

[15] J. G. Boyd and D. C. Lagoudas, "A thermodynamical constitutive model for shape memory materials. part i. the monolithic shape memory alloy," *International Journal of Plasticity*, vol. 12, no. 6, pp. 805–842, 1996.

[16] J. G. Boyd and D. C. Lagoudas, "A thermodynamical constitutive model for shape memory materials. part ii. the sma composite material," *International Journal of Plasticity*, vol. 12, no. 7, pp. 843–873, 1996.

[17] D. Lagoudas, D. Hartl, Y. Chemisky, L. Machado, and P. Popov, "Constitutive model for the numerical analysis of phase transformation in polycrystalline shape memory alloys," *International Journal of Plasticity*, vol. 32, pp. 155–183, 2012.

[18] P. S. Wellman, W. J. Peine, G. Favalora, and R. D. Howe, "Mechanical design and control of a high-bandwidth shape memory alloy tactile display," in *Experimental robotics V*, pp. 56–66, Springer, 1998.

[19] R. Gorbet, K. Morris, and R. Chau, "Mechanism of bandwidth improvement in passively cooled sma position actuators," *Smart materials and structures*, vol. 18, no. 9, p. 095013, 2009.

[20] C. Chluba, W. Ge, R. L. de Miranda, J. Strobel, L. Kienle, E. Quandt, and M. Wuttig, "Ultralow-fatigue shape memory alloy films," *Science*, vol. 348, no. 6238, pp. 1004–1007, 2015.

[21] J. Ye and Y. Gao, "Metallurgical characterization of m-wire nickel-titanium shape memory alloy used for endodontic rotary instruments during low-cycle fatigue," *Journal of endodontics*, vol. 38, no. 1, pp. 105–107, 2012.

[22] X.-L. Meng, Y.-F. Zheng, Z. Wang, and L. Zhao, "Shape memory properties of the ti36ni49hf15 high temperature shape memory alloy," *Materials Letters*, vol. 45, no. 2, pp. 128–132, 2000.

[23] P. J. S. Buenconsejo, H. Y. Kim, H. Hosoda, and S. Miyazaki, "Shape memory behavior of ti–ta and its potential as a high-temperature shape memory alloy," *Acta Materialia*, vol. 57, no. 4, pp. 1068–1077, 2009.

[24] H. Karaca, S. Saghaian, G. Ded, H. Tobe, B. Basaran, H. Maier, R. Noebe, and Y. Chumlyakov, "Effects of nanoprecipitation on the shape memory and material properties of an ni-rich nitihf high temperature shape memory alloy," *Acta Materialia*, vol. 61, no. 19, pp. 7422–7431, 2013.

[25] C. Lauhoff, N. Sommer, M. Vollmer, G. Mienert, P. Krooß, S. Böhm, and T. Niendorf, "Excellent superelasticity in a co-ni-ga high-temperature shape memory alloy processed by directed energy deposition," *Materials Research Letters*, vol. 8, no. 8, pp. 314–320, 2020.

[26] R. Featherstone and Y. H. Teh, "Improving the speed of shape memory alloy actuators by faster electrical heating," in *Experimental robotics IX*, pp. 67–76, Springer, 2006.

[27] S.-H. Song, J.-Y. Lee, H. Rodrigue, I.-S. Choi, Y. J. Kang, and S.-H. Ahn, "35 hz shape memory alloy actuator with bending-twisting mode," *Scientific reports*, vol. 6, p. 21118, 2016.

[28] H.-T. Lee, M.-S. Kim, G.-Y. Lee, C.-S. Kim, and S.-H. Ahn, "Shape memory alloy (sma)-based microscale actuators with 60% deformation rate and 1.6 khz actuation speed," *Small*, vol. 14, no. 23, p. 1801023, 2018.

[29] P. Motzki, T. Gorges, M. Kappel, M. Schmidt, G. Rizzello, and S. Seelecke, "High-speed and high-efficiency shape memory alloy actuation," *Smart Materials and Structures*, vol. 27, no. 7, p. 075047, 2018.

[30] O. Sigmund, "Design of multiphysics actuators using topology optimization–part i: One-material structures," *Computer methods in applied mechanics and engineering*, vol. 190, no. 49-50, pp. 6577–6604, 2001.

[31] O. Sigmund, "Design of multiphysics actuators using topology optimization–part ii: Two-material structures," *Computer methods in applied mechanics and engineering*, vol. 190, no. 49-50, pp. 6605–6627, 2001.

[32] Y. Li, K. Saitou, and N. Kikuchi, "Topology optimization of thermally actuated compliant mechanisms considering time-transient effect," *Finite elements in analysis and design*, vol. 40, no. 11, pp. 1317–1331, 2004.

[33] S. Cho and J.-Y. Choi, "Efficient topology optimization of thermo-elasticity problems using coupled field adjoint sensitivity analysis method," *Finite Elements in Analysis and Design*, vol. 41, no. 15, pp. 1481–1495, 2005.

[34] K. Fuchi, T. H. Ware, P. R. Buskohl, G. W. Reich, R. A. Vaia, T. J. White, and J. J. Joo, "Topology optimization for the design of folding liquid crystal elastomer actuators," *Soft matter*, vol. 11, no. 37, pp. 7288–7295, 2015.

[35] E. Jebellat, M. Baniassadi, A. Moshki, K. Wang, and M. Baghani, "Numerical investigation of smart auxetic three-dimensional meta-structures based on shape memory polymers via topology optimization," *Journal of Intelligent Material Systems and Structures*, vol. 31, no. 15, pp. 1838–1852, 2020.

[36] A. Bhattacharyya and K. A. James, "Topology optimization of shape memory polymer structures with programmable morphology," *Structural and Multidisciplinary Optimization*, vol. 63, no. 4, pp. 1863–1887, 2021.

[37] M. Langelaar, G. H. Yoon, Y. Kim, and F. Van Keulen, "Topology optimization of planar shape memory alloy thermal actuators using element connectivity parameterization," *International journal for numerical methods in engineering*, vol. 88, no. 9, pp. 817–840, 2011.

[38] Z. Kang and K. A. James, "Thermomechanical topology optimization of shape-memory alloy structures using a transient bilevel adjoint method," *International Journal for Numerical Methods in Engineering*, vol. 121, no. 11, pp. 2558–2580, 2020.

[39] C. Kittel and H. Kroemer, "Thermal physics," 1998.

[40] M. Qidwai and D. Lagoudas, "Numerical implementation of a shape memory alloy thermo-mechanical constitutive model using return mapping algorithms," *International Journal for Numerical Methods in Engineering*, vol. 47, no. 6, pp. 1123–1168, 2000.

[41] M. Qidwai and D. Lagoudas, "On thermomechanics and transformation surfaces of polycrystalline niti shape memory alloy material," *International journal of plasticity*, vol. 16, no. 10-11, pp. 1309–1343, 2000.

[42] J. C. Simo and T. J. Hughes, *Computational inelasticity*, vol. 7. Springer Science & Business Media, 2006.

[43] Z. Kang, D. A. Tortorelli, and K. A. James, "Parallel projection-a new return mapping method for finite element modeling of shape memory alloys," *ArXiv*, 2021.

[44] M. P. Bendsøe and O. Sigmund, "Material interpolation schemes in topology optimization," *Archive of applied mechanics*, vol. 69, no. 9-10, pp. 635–654, 1999.

[45] K. A. James, "Multiphase topology design with optimal material selection using an inverse p-norm function," *International Journal for Numerical Methods in Engineering*, vol. 114, no. 9, pp. 999–1017, 2018.

[46] Z. Kang and K. A. James, "Multimaterial topology design for optimal elastic and thermal response with material-specific temperature constraints," *International Journal for Numerical Methods in Engineering*, vol. 117, no. 10, pp. 1019–1037, 2019.

[47] K. Svanberg, "The method of moving asymptotes—a new method for structural optimization," *International journal for numerical methods in engineering*, vol. 24, no. 2, pp. 359–373, 1987.

[48] W. D. Callister and D. G. Rethwisch, *Materials science and engineering: an introduction*, vol. 9. Wiley New York, 2018.

[49] R. Tian, C. Hang, Y. Tian, and J. Feng, "Brittle fracture induced by phase transformation of ni-cu-sn intermetallic compounds in sn-3ag-0.5 cu/ni solder joints under extreme temperature environment," *Journal of Alloys and Compounds*, vol. 777, pp. 463–471, 2019.

[50] C. Buchanan and L. Gardner, "Metal 3D printing in construction: A review of methods, research, applications, opportunities and challenges," *Engineering Structures*, vol. 180, pp. 332–348, 2019.

[51] T. D. Ngo, A. Kashani, G. Imbalzano, K. T. Nguyen, and D. Hui, "Additive manufacturing (3d printing): A review of materials, methods, applications and challenges," *Composites Part B: Engineering*, vol. 143, pp. 172–196, 2018.

[52] F. Zhang, *The Schur complement and its applications*, vol. 4. Springer Science & Business Media, 2006.